

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2000, 08:52:43 ; Search time 374.18 Seconds
(without alignments)
10074.554 Million cell updates/sec

Title: US-08-852-495C-1_COPY_140000_169000
Perfect score: 29001
Sequence: 1 CCTCCCAATCCCATATGAC.....TGTATCAGAGAAATTACC 29001

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCUTS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	76.8	0.3	105	US-08-481-658B-65	Sequence 65, Appl
C 2	76.8	0.3	105	US-08-477-504A-65	Sequence 65, Appl
C 3	76.8	0.3	105	US-08-486-756A-65	Sequence 65, Appl
C 4	76.8	0.3	105	US-08-485-862B-65	Sequence 65, Appl
C 5	76.8	0.3	105	US-08-787-739-65	Sequence 65, Appl
C 6	71.2	0.2	105	US-08-481-658B-65	Sequence 65, Appl
C 7	71.2	0.2	105	US-08-477-504A-65	Sequence 65, Appl
C 8	71.2	0.2	105	US-08-486-756A-65	Sequence 65, Appl
C 9	71.2	0.2	105	US-08-485-862B-65	Sequence 65, Appl
C 10	71.2	0.2	105	US-08-787-739-65	Sequence 65, Appl
C 11	65.4	0.2	84	US-08-454-557C-91	Sequence 91, Appl
C 12	65.4	0.2	84	US-08-340-426D-91	Sequence 91, Appl
C 13	65.4	0.2	84	US-08-450-673C-91	Sequence 91, Appl
C 14	65.4	0.2	84	PCR-US95-17111A-91	Sequence 91, Appl
C 15	60.4	0.2	78	US-08-454-557C-70	Sequence 70, Appl
C 16	60.4	0.2	78	US-08-340-426D-70	Sequence 70, Appl
C 17	60.4	0.2	78	US-08-450-673C-70	Sequence 70, Appl
C 18	60.4	0.2	78	PCR-US95-17111A-70	Sequence 70, Appl
C 19	57	0.2	76	US-08-454-557C-69	Sequence 69, Appl
C 20	57	0.2	76	US-08-340-426D-69	Sequence 69, Appl
C 21	57	0.2	76	US-08-450-673C-69	Sequence 69, Appl
C 22	57	0.2	76	PCR-US95-17111A-69	Sequence 69, Appl
C 23	56	0.2	85	US-08-454-557C-92	Sequence 92, Appl
C 24	56	0.2	85	US-08-340-426D-92	Sequence 92, Appl
C 25	56	0.2	85	US-08-450-673C-92	Sequence 92, Appl
C 26	56	0.2	85	PCR-US95-17111A-92	Sequence 92, Appl
C 27	55.6	0.2	78	US-08-454-557C-70	Sequence 70, Appl

C 28	55.6	0.2	78	US-08-340-426D-70	Sequence 70, Appl
C 29	55.6	0.2	78	US-08-450-673C-70	Sequence 70, Appl
C 30	55.6	0.2	78	PCR-US95-17111A-70	Sequence 70, Appl
C 31	55.2	0.2	60	US-08-454-557C-57	Sequence 57, Appl
C 32	55.2	0.2	60	US-08-340-426D-57	Sequence 57, Appl
C 33	55.2	0.2	60	US-08-450-673C-57	Sequence 57, Appl
C 34	55.2	0.2	60	PCR-US95-17111A-57	Sequence 57, Appl
C 35	55	0.2	84	US-08-454-557C-91	Sequence 91, Appl
C 36	55	0.2	84	US-08-340-426D-91	Sequence 91, Appl
C 37	55	0.2	84	US-08-450-673C-91	Sequence 91, Appl
C 38	55	0.2	84	PCR-US95-17111A-91	Sequence 91, Appl
C 39	53.2	0.2	83	US-08-481-658B-66	Sequence 66, Appl
C 40	53.2	0.2	83	US-08-477-504A-66	Sequence 66, Appl
C 41	53.2	0.2	83	US-08-486-756A-66	Sequence 66, Appl
C 42	53.2	0.2	83	US-08-485-862B-66	Sequence 66, Appl
C 43	53.2	0.2	83	US-08-787-739-66	Sequence 66, Appl
C 44	50	0.2	76	US-08-454-557C-69	Sequence 69, Appl
C 45	50	0.2	76	US-08-340-426D-69	Sequence 69, Appl

ALIGNMENTS

RESULT 1
US-08-481-658B-65/C
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (Epo)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-0727
; TELEFAX: 415-435-2034
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-481-658B-65
Query Match 0.3%; Score 76.8; DB 4; Length 105;
Best Local Similarity 83.7%; Pred. No. 4e-07;

Matches 87; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 24042 ATCCGACACTTTGGAGGCTGAGGCTGAATCAGCAGGTCAAGATCAAGACCATC 24101

DB 105 ATCCGACACTTTGGAGGCTGAGGCTGAATCAGCAGGTCAAGATCAAGATTTAGAGCAGC 46

QY 24102 CTGGCCAAATGTGTGAACCCCTGCTCTACTAAATACAAAAA 24145

DB 45 CTGGCCAAATGTGTGAACCCCTGCTCTACTAAAGATGTAAAAA 2

RESULT 2

US-08-477-504A-65/C

; Sequence 65, Application US/08477504A

; Patent No. 5972353

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/477,504A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3D

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-477-504A-65

Query Match 0.3%; Score 76.8; DB 4; Length 105;

Best Local Similarity 83.7%; Pred. No. 4e-07;

Matches 87; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 24042 ATCCGACACTTTGGAGGCTGAGGCTGAATCAGCAGGTCAAGATCAAGACCATC 24101

DB 105 ATCCGACACTTTGGAGGCTGAGGCTGAATCAGCAGGTCAAGATTTAGAGCAGC 46

QY 24102 CTGGCCAAATGTGTGAACCCCTGCTCTACTAAATACAAAAA 24145

DB 45 CTGGCCAAATGTGTGAACCCCTGCTCTACTAAAGATGTAAAAA 2

RESULT 3

US-08-486-756A-65/C

; Sequence 65, Application US/08486756A

; Patent No. 5981711

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,756A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-486-756A-65

Query Match 0.3%; Score 76.8; DB 4; Length 105;

Best Local Similarity 83.7%; Pred. No. 4e-07; Mismatches 17; Indels 0; Gaps 0;

Matches 87; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 24042 ATCCGACACTTTGGAGGCTGAGGCTGAATCAGCAGGTCAAGATCAAGACCATC 24101

DB 105 ATCCGACACTTTGGAGGCTGAGGCTGAATCAGCAGGTCAAGATTTAGAGCAGC 46

QY 24102 CTGGCCAAATGTGTGAACCCCTGCTCTACTAAATACAAAAA 24145

DB 45 CTGGCCAAATGTGTGAACCCCTGCTCTACTAAAGATGTAAAAA 2

RESULT 4

US-08-485-862B-65/C

; Sequence 65, Application US/08485862B

; Patent No. 5989838

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 08/485,862B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ. ID NO.: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-485-862B-65

Query Match 0.3%; Score 76.8; DB 4; Length 105;
Best Local Similarity 83.7%; Pred. No. 4e-07;
Matches 87; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 24042 ATCCGACACTTTGGGAGCGTGAGTGATCAGAGGTGAGATCAAGACCATC 24101
|||||
DB 105 ATCCGACACTTTGGGAGCGTGAGTGATCAGAGGTGAGATCAAGGTGAGAGCAGC 46
QY 24102 CTGGCCAACTGTTGAAACCCGCTCTCTACTAAATACAAAAA 24145
|||||
DB 45 CTGGCCAACTGTTGAAACCCGCTCTCTACTAAAGATGTAATAA 2

RESULT 5
US-08-787-739-65/c
Sequence 65, Application US/08787739
Patent No. 6027887
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 96
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 369 Pine Street, Suite 610
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/787,739
FILING DATE: 24-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,049
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/486,756
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/481,658
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/485,862
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,863
FILING DATE: 07-JUN-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 08/487,077
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-981-0332
TELEFAX: 415-981-2034
INFORMATION FOR SEQ. ID NO.: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-787-739-65

Query Match 0.3%; Score 76.8; DB 5; Length 105;
Best Local Similarity 83.7%; Pred. No. 4e-07;
Matches 87; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 24042 ATCCGACACTTTGGGAGCGTGAGTGATCAGAGGTGAGATCAAGACCATC 24101
|||||
DB 105 ATCCGACACTTTGGGAGCGTGAGTGATCAGAGGTGAGATCAAGGTGAGAGCAGC 46
QY 24102 CTGGCCAACTGTTGAAACCCGCTCTCTACTAAATACAAAAA 24145
|||||
DB 45 CTGGCCAACTGTTGAAACCCGCTCTCTACTAAAGATGTAATAA 2

RESULT 6
US-08-481-658B-65
Sequence 65, Application US/08481658B
Patent No. 5955075
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,658B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.2%; Score 71.2; DB 4; Length 105;
Best Local Similarity 85.8%; Pred. No. 5.4e-06;
Matches 91; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

QY 16710 TTTTGTATTTTATAGTAGAGATAGGCTTCACAAATGCTGGCCAGGCTGCTCAAACTCC 16769
DB 2 TTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCAAACTCC 61
QY 16770 TGCCCTCAAGTGATCCTCCTGCTGCGCTGCCCAATGCTGCTGGAT 16815
DB 62 TGACCT--TGTGATCCACACAGCCTGCGCTCCCAAGTGTGCGAT 105

RESULT 7
US-08-477-504A-65
Sequence 65, Application US/08477504A
Patent No. 5972353
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,504A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:

NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-477-504A-65

Query Match 0.2%; Score 71.2; DB 4; Length 105;
Best Local Similarity 85.8%; Pred. No. 5.4e-06;
Matches 91; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

QY 16710 TTTTGTATTTTATAGTAGAGATAGGCTTCACAAATGCTGGCCAGGCTGCTCAAACTCC 16769
DB 2 TTTTACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCAAACTCC 61
QY 16770 TGCCCTCAAGTGATCCTCCTGCTGCGCTGCCCAATGCTGCTGGAT 16815
DB 62 TGACCT--TGTGATCCACACAGCCTGCGCTCCCAAGTGTGCGAT 105

RESULT 8
US-08-486-756A-65
Sequence 65, Application US/08486756A
Patent No. 5981711
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,756A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

ANTI-SENSE
;
ME-09-797-739-6


```

: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/450,673C
: FILING DATE: 30-MAY-1995
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: LUDWIG, Steven R.
: REGISTRATION NUMBER: 36, 203
: REFERENCE/DOCKET NUMBER: 0609, 3840004
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 371-2600
: TELEFAX: (202) 371-2540
: INFORMATION FOR SEQ ID NO: 91:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 84 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: both
: US-08-450-673C-91

```

	Query Match	Similarity	Best Local	Conservative	Matches	Score	DB	Length	Indels	Gaps
		0.2%	86.7%	0	72	NO. 7.4e-05	4	84	0	0
QY	24032	CACGCTTAAATCCACAGACTTTGGAGAGCTAGAGTGGTGAATACAGAGTGCAGAGAT					24091			
DB	83	CACGCTTGAATGCCAGCACTTTGGAGAGCTTAGGGGGGGGAGTACGAGTCTCAGAGTT				24				
QY	24092	CAGACCATCTGGCCAAACATGG				24114				
DB	23	CGACACGAGCCGTGATGAACATGG				1				

```

1      RESULT 14
2      PCT-US95-17111A-91/C
3      : Sequence 91, Application PC/TUS9517111A
4      : GENERAL INFORMATION:
5      :   APPLICANT: de la Monte, Suzanne
6      :   APPLICANT: Wands, Jack R.
7      :   TITLE OF INVENTION: Neural Thread Protein Gene Expression and
8      :   TITLE OF INVENTION: Detection of Alzheimer's Disease
9      :   NUMBER OF SEQUENCES: 121
10     :   CORRESPONDENCE ADDRESSES:
11     :     ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
12     :     STREET: 1100 New York Avenue, Suite 600
13     :     CITY: Washington
14     :     STATE: D.C.
15     :     COUNTRY: U.S.A.
16     :     ZIP: 20005-3934
17     :   COMPUTER READABLE FORM:
18     :     MEDIUM TYPE: Floppy disk
19     :     COMPUTER: IBM PC compatible
20     :     OPERATING SYSTEM: PC-DOS/MS-DOS
21     :     SOFTWARE: Patent In Release #1.0, Version #1.25
22     :   CURRENT APPLICATION DATA:
23     :     APPLICATION NUMBER: PCT/US95/17111A
24     :     FILING DATE:
25     :     CLASSIFICATION:
26     :     PRIOR APPLICATION DATA:
27     :       APPLICATION NUMBER: 08/340,426
28     :       FILING DATE: 14-NOV-1994
29     :     ATTORNEY/AGENT INFORMATION:
30     :       NAME: Ludwig, Steven R.
31     :       REGISTRATION NUMBER: 36,203
32     :       REFERENCE/DOCKET NUMBER: 0609.3840002
33     :     TELECOMMUNICATION INFORMATION:
34     :       TELEPHONE: (202) 371-2600
35     :       TELEFAX: (202) 371-2540
36     :     INFORMATION FOR SEQ ID NO: 91:
37     :       SEQUENCE CHARACTERISTICS:
38     :         LENGTH: 84 base pairs
39     :         TYPE: nucleic acid
40     :         STRANDEDNESS: both

```

TOPOLOGY: both
PCT-US95-17111A-91

Query Match	0.28;	Score 65.4;	DB 6;	Length 84;
Best Local Similarity	86.78;	Pred. No. 7.4e-05;		
Matches	72;	Conservative	0;	Mismatches 11;
			Indels	0;
			Gaps	0;

Accession	Gene	Protein	Accession	Gene	Protein
Oy 24032	CACGCGCTTAAATCCAGGACTTTGGAGAGCTT	AGTGTGAATCAGAGGTCCAGAGAT	24091		
Db 83	CACGCGTCTTAATCCAGGACTTTGGAGAGCTT	AGGCGGSCGATCATCAGAGTCCAGAGATT	24		
Oy 24092	CAGAGCATCCTGGGCCAATG	24114			
Db 23	CGACACGACCGATGAATCAATGG	1			

RESULT 15
 US-08-454-557C-70
 Sequence 70: Application US/08454557C
 Patent No. 5830670
 GENERAL INFORMATION:
 APPLICANT: de la Monte, Suzanne
 APPLICANT: Wands, Jack R.
 TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
 TITLE OF INVENTION: Of Alzheimer's Disease
 NUMBER OF SEQUENCES: 121
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
 STREET: 1100 New York Avenue, Suite 600
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005-3934
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/454,557C
 FILING DATE: 30-MAY-1995
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Ludwig, Steven R.
 REGISTRATION NUMBER: 36,203
 REFERENCE/DOCKET NUMBER: 0609, 3840003
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 70:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 78 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: both

	Query Match	0.2%	Score 60.4;	DB 3;	Length 78;	
	Best Local Similarity	85.9%;	Pred. No. 0.00074;			
	Matches	67;	Conservative	0;	Mismatches 11;	Indels 0; Gaps 0.
QY	9486	ACCATGCCCCGGTATTTTGTAATTTTTAGTAGAGACAGCGTTTCACCATTCTGTGGCCAGG	9545			
Dd	1	ACAACGCCACAGTAAATAATTTGTAATTTTTAGTAGAGATGGGTTTCTCATTGTTATCAGG	60			
QY	9546	CTGGGTGTGAACCTCCTGA	9563			
Dd	61	CTGGTGTGAACCTCCTGA	78			

Wed Jun 21 14:43:23 2000

us-08-852-495c-1_copy_140000_169000.rni

Page 8

Search completed: June 16, 2000, 20:15:34
Job time: 199264 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2000, 20:09:08 ; Search time 29137.4 Seconds
(without alignments)
-968.237 Million cell updates/sec

Title: US-08-852-495C-1_COPY_168000_197000
Perfect score: 29001
Sequence: 1 TGTATAGAGGAAAAAAGCA.....TAGATAAAGCTGTCTCTT 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_cm:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl1:*
10: gb_pl2:*
11: gb_pl3:*
12: gb_ro:*
13: gb_sy:*
14: gb_sy:*
15: gb_un:*
16: gb_v1:*
17: em_fun:*
18: em_hum1:*
19: em_hum2:*
20: em_hum3:*
21: em_hum4:*
22: em_hum5:*
23: em_hum6:*
24: em_hum7:*
25: em_hum8:*
26: em_hum9:*
27: em_hum10:*
28: em_hum11:*
29: em_hum12:*
30: em_hum13:*
31: em_hum14:*
32: em_hum15:*
33: em_hum16:*
34: em_hum17:*
35: em_hum18:*
36: em_hum19:*
37: em_hum20:*
38: em_hum21:*
39: em_hum22:*
40: em_hum23:*
41: em_hum24:*
42: em_hum25:*
43: em_hum26:*
44: em_hum27:*

45: gb_hcg7:*
46: em_hcg1:*
47: em_hcg2:*
48: em_hcg3:*
49: em_hcg4:*
50: em_hcg5:*
51: gb_pl3:*
52: gb_pl4:*
53: gb_hcg8:*
54: gb_hcg9:*
55: gb_hcg10:*
56: gb_hcg11:*
57: gb_hcg12:*
58: gb_hcg13:*
59: gb_hcg14:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	95.2	0.3	108	HSIDLBN2	X05250 Human LDL-r
2	92	0.3	108	HSIDLBN2	X05250 Human LDL-r
3	85	0.3	107	HUMALCE162	M87924 Human LDL-r
4	84.6	0.3	108	HSIDLBN1	X05249 Human LDL-r
5	84.6	0.3	108	HSIDLBN2	X05251 Human LDL-r
6	83	0.3	108	HSIDLBN1	X05249 Human LDL-r
7	83	0.3	108	HSIDLBN2	X05251 Human LDL-r
8	82.2	0.3	103	HUMALCE221	M87896 Human LDL-r
9	81	0.3	108	HSIDLBN1	X05249 Human LDL-r
10	81	0.3	108	HSIDLBN2	X05251 Human LDL-r
11	79.2	0.3	103	HSIDLBN1	X05249 Human LDL-r
12	78	0.3	107	HUMALCE162	M87924 Human LDL-r
13	77.6	0.3	108	HUMALCE221	M87896 Human LDL-r
14	77.2	0.3	110	HSIDLBN1	X05249 Human LDL-r
15	76.6	0.3	108	HSIDLBN2	X05251 Human LDL-r
16	76.6	0.3	110	HSIDLBN1	X05249 Human LDL-r
17	75.2	0.3	97	HUMALCE221	M87896 Human LDL-r
18	74.4	0.3	103	HSIDLBN1	X05249 Human LDL-r
19	74.4	0.3	104	HUMALCE221	M87896 Human LDL-r
20	74.6	0.3	108	HSIDLBN1	X05249 Human LDL-r
21	73.4	0.3	91	HSIDLBN1	X05249 Human LDL-r
22	72.4	0.2	108	HSIDLBN1	X05249 Human LDL-r
23	72.4	0.2	108	HSIDLBN1	X05249 Human LDL-r
24	72.4	0.2	110	HSIDLBN1	X05249 Human LDL-r
25	72	0.2	97	HUMALCE221	M87896 Human LDL-r
26	72.2	0.2	108	HSIDLBN1	X05249 Human LDL-r
27	71.4	0.2	107	HSIDLBN1	X05249 Human LDL-r
28	71.4	0.2	108	HSIDLBN1	X05249 Human LDL-r
29	70.8	0.2	90	HSIDLBN1	X05249 Human LDL-r
30	70.8	0.2	100	HSIDLBN1	X05249 Human LDL-r
31	70.2	0.2	97	HSIDLBN1	X05249 Human LDL-r
32	70.4	0.2	106	HSIDLBN1	X05249 Human LDL-r
33	69.4	0.2	108	HSIDLBN1	X05249 Human LDL-r
34	68.6	0.2	84	HSIDLBN1	X05249 Human LDL-r
35	68.4	0.2	95	HSIDLBN1	X05249 Human LDL-r
36	68.2	0.2	99	HSIDLBN1	X05249 Human LDL-r
37	67.8	0.2	108	HSIDLBN1	X05249 Human LDL-r
38	67.4	0.2	79	HSIDLBN1	X05249 Human LDL-r
39	67.4	0.2	102	HSIDLBN1	X05249 Human LDL-r
40	66.8	0.2	95	HSIDLBN1	X05249 Human LDL-r
41	66.8	0.2	97	HSIDLBN1	X05249 Human LDL-r
42	66	0.2	97	HSIDLBN1	X05249 Human LDL-r
43	65.8	0.2	106	HSIDLBN1	X05249 Human LDL-r
44	64.8	0.2	79	HSIDLBN1	X05249 Human LDL-r
45	65	0.2	99	HSIDLBN1	X05249 Human LDL-r

ALIGNMENTS

	RESULT	1
HSLIDLNRN/c	LOCUS	
DEFINITION	Human LDL-receptor gene intron 14 fragment (normal gene).	PRI
ACCESSION	X05250	20-MAY-1992
VERSION	X05250.1 GI:34337	
KEYWORDS	Alu repetitive sequence; Low density lipoprotein receptor.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eunayrola, Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;	
AUTHORS	Primates; Catarrhini; Hominiidae; Homo.	
TITLE	1 (bases 1 to 108)	
JOURNAL MEDLINE	Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,	
COMMENT	Williamson,R. and Humphries,S. Unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia <i>Eur. J. Biochem.</i> 164 (1), 77-81 (1987) 87161901	
FEATURES	See X05252 for deletion junction	
SOURCE	Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.	
	Location/Qualifiers	
	1..108	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	1..108	
	/note="Intron XIV fragment"	
BASE COUNT	28 a 23 c 39 g 18 t	
ORIGIN		

	Query Match	Similarity	0.3%	Score 95.2	DB 10	Length 108
	Best Local	Similarity	92.6%	Pred. No. 3e-06		
	Matches 100	Conservative	0	Mismatches 8	Indels 0	Gaps 0
Oy	22680	CTTGGCTCACTGCACACCTCCGCTCTGGGTTCAAGCAATTCGCATGCTCAGCCCTCCG	22739			
Db	108	CTCGGCTCACTGCACACCTTCGCTCTGGGTTCAAGCAATTCCTCGCTCAGCCCTCCG	49			
Oy	22740	AGTACTGGGATTACAGGCACATACACATGACATGGCAATTTTGT	22787			
Db	48	AGTACTGGGATTACAGGCACATTCGACACACCTTCGGCAATTTTGT	1			

RESULT	2	
HSIDLNR2		
LOCUS	108 bp	DNA
DEFINITION	Human LDL-receptor gene intron 14 fragment (normal gene).	PRI
ACCESSION	X05250	
VERSION	X05250.1 GI:34337	
KEYWORDS	Alu repetitive sequence; low density lipoprotein receptor.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Euthera;	
AUTHORS	Primates; Catarrhini; Homnidae; Homo.	
	1 (bases 1 to 108)	
	Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,	
	Williamson,R. and Humphries,S.	
TITLE	Unequal crossing-over between two alu-repetitive DNA sequences in	
	the low-density lipoprotein-receptor gene. A possible mechanism for	
JOURNAL	the defect in a patient with familial hypercholesterolaemia	
MEDLINE	Eur. J. Biochem. 164 (1), 77-81 (1987)	
COMMENT	87161901	
	See X05252 for deletion junction	
FEATURES	Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.	
source	Location/Qualifiers	
	1..108	

Feature	Start	End	Score	Strand	Frame	Phase	Gene
Intron	1	108					
Exon	109	228					
Intron	229	390					
Exon	391	552					
Intron	553	714					
Exon	715	876					
Intron	877	1038					
Exon	1039	1200					
Intron	1201	1362					
Exon	1363	1524					
Intron	1525	1686					
Exon	1687	1848					
Intron	1849	2010					
Exon	2011	2172					
Intron	2173	2334					
Exon	2335	2496					
Intron	2497	2658					
Exon	2659	2820					
Intron	2821	2982					
Exon	2983	3144					
Intron	3145	3306					
Exon	3307	3468					
Intron	3469	3630					
Exon	3631	3792					
Intron	3793	3954					
Exon	3955	4116					
Intron	4117	4278					
Exon	4279	4440					
Intron	4441	4602					
Exon	4603	4764					
Intron	4765	4926					
Exon	4927	5088					
Intron	5089	5250					
Exon	5251	5412					
Intron	5413	5574					
Exon	5575	5736					
Intron	5737	5898					
Exon	5899	6060					
Intron	6061	6222					
Exon	6223	6384					
Intron	6385	6546					
Exon	6547	6708					
Intron	6709	6870					
Exon	6871	7032					
Intron	7033	7194					
Exon	7195	7356					
Intron	7357	7518					
Exon	7519	7680					
Intron	7681	7842					
Exon	7843	8004					
Intron	8005	8166					
Exon	8167	8328					
Intron	8329	8490					
Exon	8491	8652					
Intron	8653	8814					
Exon	8815	8976					
Intron	8977	9138					
Exon	9139	9300					
Intron	9301	9462					
Exon	9463	9624					
Intron	9625	9786					
Exon	9787	9948					
Intron	9949	10110					
Exon	10111	10272					
Intron	10273	10434					
Exon	10435	10596					
Intron	10597	10758					
Exon	10759	10920					

ORIGIN

Query Match	0.38;	Score 92;	DB 10;	Length 108;
Best Local Similarity	90.7%;	Pred. No. 1e-05;		
Matches 98;	Conservative	0;	Mismatches 10;	Indels 0;
				Gaps 0;

```
Qy 16843 ACAAAATTTAGCCAGGCGCTGTGGCATCTGCCCTTAGTCCACGACTACTGGGACACTGAG 16902Z
      |||||
Db   1 ACAAATAATTAGCAGGCGCTGTGGCAGGTCCTGAATCCACGACTACTGGGAGCGTGAG 60
```

```

Oy 16903 GCAGGAAATCACTTGAAACCCAGGAGGAGAGATTGCAGTGAGCTGAG 169
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 61 GCAGGAAATTGCTTGAAACCCAGGAGGAGAGGTTGCAGTGAGCCGAG 108

```

RESULT 3
HUMALCE162

LOCUS	HUMALCE162	107 bp ss-RNA	PRI
DEFINITION	Human carcinoma cell-derived Alu RNA transcript, clone CE162		
ACCESSION	M87924		
VERSION	M87924.1	GI:174871	

SOURCE ORGANISM	Embryo	carcinoma	CDNA	to other RNA.
Homo sapiens				
Homo sapiens				

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
1 (bases 1 to 107)
Smettett, D., Richer, C., Deragon, J.-M. and Labuda, D.
Alu RNA transcripts in human embryonal carcinoma cells. Model of
post-transcriptional selection of master sequences
Nucleic Acids Res. 1993; 21: 1003-1010

Source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Ntera3D1"
/dev_stage="Embryo"
/sex="male"
/tissue_type="carcinoma"
BASE_COUNT      28 a      30 c      35 g      14 t
ORIGIN

```

Query Match	0.3%;	Score 85;	DB 9;	Length 107;
Best Local Similarity	90.1%;	Pred. No. 0.00014;		
Matches	91;	Conservative	0;	Mismatches 10;
			Indels	0;
			Gaps	0;

```

Oy 16902 GCGAGGAGATCACTTGAACCCAGGAGCGAGAGATTGCAGCTGAGATCGGCCACT 16961
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   5 GCGAGAGAGATTGGCGTGAACCCGGAGCGGAGCTTGCAGTGAAGCCGAGATCGGCCACT 64

```

```
Oy   16962 GCATTCAGCCTGGGACACAGAAGCACTCGTGTCAA AAA    17002  
      ||| | | | | | | | | | | | | | | | |  
Db     65 GCACTCAGCCCTGGGCACAGAGCGAACA TCGTGTA AA A    105
```

RESULT	4
HSLDLRD1	
LOCUS	HSLDLRD1 108 bp DNA PRI 20-MAY-1992
DEFINITION	Human LDL-receptor mutated gene with intron 2 deletion junction
ACCESSION	X05249
VERSION	X05249.1 GI:34335
KEYWORDS	All repetitive sequence; low density lipoprotein receptor.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 108)	Horsthemke, B., Beisiegel, U., Dunning, A., Havinga, J.R., Williamson, R. and Humphries, S.	Unequal crossing-over between two α -globin repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for

AUTHORS Horsthemke, B., Beisiegel, U., Dunning, A., Havinga, J.R., Williamson, R. and Humphries, S.
TITLE Unequal crossing-over between two *alu*-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia
JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901
COMMENT *source: hypercholesterol aemia
See X05250 for corresponding normal gene sequence
In the defective LDL-receptor gene the deletion occurred between two *alu*-repetitive sequences, that are in the same direction, the deletion eliminates exons 13 and 14 and changes the reading frame of the resulting spliced mRNA.
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.

FEATURES
source
1..108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="blood leukocytes from a patient with familial"
1..108
/note="intron XIV fragment"
BASE COUNT 28 a 20 c 40 g 20 t
ORIGIN

Query Match
Best Local Similarity 0.3%; Score 83; DB 10; Length 108;
Matches 92; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

OY 16843 ACAAAATTAGGCGAGCGCTGGCATCTGCTGTATGCCAGCTACTCGGAGCACTGAG 16902
DB 1 ACAAAATTAGGCGAGCGCTGGCATCTGCTGTATGCCAGCTACTCGGAGCGCTGAG 60
OY 16903 GCAGAGAAATCTGTAACCCGAGGAGGAGATGTCAGTGCAGTGCAGTGA 16949
DB 61 GCAGGAAATGCTTGAACCCGAGGAGGAGGAGTGTGTGAGGCGA 107

RESULT 8
HUMALCE221 103 bp ss-RNA PRI 15-APR-1994
LOCUS Human carcinoma cell-derived *Alu* RNA transcript, clone CE221.
DEFINITION M87896
ACCESSION M87896.1 GI:174874
VERSION
KEYWORDS *Alu* repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE Smet, D., Richey, C., Deragon, J.-M. and Labuda, D.
TITLE *Alu* RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
source
1..103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="NTera2D1"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"
BASE COUNT 25 a 27 c 33 g 18 t
ORIGIN

Query Match
Best Local Similarity 0.3%; Score 82.2; DB 9; Length 103;
Matches 90; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY 16868 AACTGCTCTAGTCCAGCTACTCGGAGACATGAGGAGAGAAATCTGGAACCCAGGA 16927
DB 1 AACTGCTCTAGTCCAGCTACTCGGAGACATGAGGAGAGAAATCTGGAACCCAGGA 60

OY 16928 GGCAGAGATTGAGTGCAGATCGGCCACATTCATTCAG 16970
DB 61 GGCAGAGATTGAGTGCAGATCGGCCACATTCATTCAG 103

RESULT 9
HSU67803 108 bp RNA PRI 01-AUG-1997
LOCUS Human small cytoplasmic *Alu* transcript.
DEFINITION U67803
ACCESSION U67803 GI:2289917
VERSION
KEYWORDS *Alu*.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE Shaikh, T.H., Roy, A.M., Kim, J., Batzer, M.A. and Deininger, P.L.
TITLE cDNAs derived from primary and small cytoplasmic *Alu* (scAlu) transcripts
JOURNAL J. Mol. Biol. 271 (2), 222-234 (1997)
MEDLINE 97415756
REFERENCE 2 (bases 1 to 108)
AUTHORS Shaikh, T.H., Kim, J., Batzer, M.A. and Deininger, P.L.
JOURNAL Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The Children's Hospital of Philadelphia, 1004F Abramson Research Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES
source
1..108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TscAlu2"
1..108
/note="scAlu"
/rpl_family="Alu"
repeat_region /rpl_type="dispersed"
BASE COUNT 23 a 39 c 30 g 16 t
ORIGIN

Query Match
Best Local Similarity 0.3%; Score 81; DB 11; Length 108;
Matches 87; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 16737 GCCTTAATCCAGACATCTGGGAGCCAGGAGGAGACATGAGGAGGACTTGA 16796
DB 1 GCCTTAATCCAGACATCTGGGAGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 60
OY 16797 GACGAGCTGACACATGATGTAACCTGTCTTAC 16833
DB 61 GACGAGCTGACACATGATGTAACCTGTCTTAC 97

RESULT 10
HSU67808 108 bp RNA PRI 01-AUG-1997
LOCUS Human small cytoplasmic *Alu* transcript.
DEFINITION U67808
ACCESSION U67808.1 GI:2289922
VERSION
KEYWORDS *Alu*.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE Shaikh, T.H., Roy, A.M., Kim, J., Batzer, M.A. and Deininger, P.L.
TITLE cDNAs derived from primary and small cytoplasmic *Alu* (scAlu) transcripts
JOURNAL J. Mol. Biol. 271 (2), 222-234 (1997)
MEDLINE 97415756
REFERENCE 2 (bases 1 to 108)

AUTHORS Shaikh, T.H., Kim, J., Batzer, M.A. and Delinger, P.L.
TITLE Direct Submission
JOURNAL Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The Children's Hospital of Philadelphia, 1004F Abramson Research Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES
source 1. 108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Tscaliu7"
repeat_region 1. 108
/note="scaliu"
/rpt_family="Alu"
/rpt_type="dispersed"
BASE COUNT 22 a 37 c 28 g 21 t
ORIGIN

Query Match 0.3%; Score 81; DB 11; Length 108;
Best Local Similarity 89.7%; Pred. No. 0.00062;
Matches 87; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 16737 GCCTGTAATCCAGCAGCTTTGGAGGCCAAGCGGAGATCAGAGTCAGAGTTTGA 16796
|||||
Db 1 GCCTGTAATCCAGCAGCTTTGGAGGCCAAGCGGAGATCAGAGTCAGAGTTTGA 60
|||||

Qy 16797 GACCAAGCTGACCAACATGTGTGAACCCGTCTCTTAC 16833
|||||
Db 61 GACCAAGCTGACCAACATGTGTGAACCCGTCTCTTCC 97
|||||

RESULT 11
HS8IC8R 103 bp DNA STS 05-SEP-1991
LOCUS HS8IC8R Human sequence tagged site 8IC8R DNA from 19q13.
ACCESSION X57789
VERSION X57789.1 GI:23938
KEYWORDS STS; myotonic dystrophy.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 103)
Aldridge, F.L.
Direct Submission
Submitted (12-FEB-1991) F.L. Aldridge, ICI Pharmaceuticals,
Alderley Park, Macclesfield, Cheshire, SK10 4TG, UK
2 (bases 1 to 103)
Butler, R., Riley, J.H., Ogilvie, D.J., Anand, R., Buxton, J.,
Davies, J., Johnson, K. and Markham, A.F.
Two sequence-tagged sites defining the ends of a 380 kb YAC clone
from 19q13
Nucleic Acids Res. 19 (17), 4787 (1991)
91367697
See also X57788 for STS 8IC8L.
COMMENT Location/Qualifiers
FEATURES
source 1. 103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19q13"
/germline
/clone_lib="YAC library: ICI"
/clone="8IC8"
BASE COUNT 29 a 28 c 23 g 22 t 1 others
ORIGIN

Query Match 0.3%; Score 79.2; DB 13; Length 103;
Best Local Similarity 86.1%; Pred. No. 0.0012;
Matches 87; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Qy 27423 TACCTTAATCCAGCAGCTTTGGAGACTGAGTGGGTGATCATCTGAGTCAGAGTT 27482
|||||

Db 102 TGCCATAATTTGNGACATTTGGAGGTGAGGTGGTGATCACTTAAGTCAGAGTT 43
Qy 27483 CAAACACAGCTGGCCACATGGTGAAGAACCATCTCTACT 27523
|||||
Db 42 CTGACACAGCTGGCCACATGGTGAAGAACCATCTCTACT 2

RESULT 12
HUMALCE162/c 107 bp ss-RNA PRI 15-APR-1994
LOCUS HUMALCE162 Human carcinoma cell-derived Alu RNA transcript, clone CE162.
ACCESSION M87924
VERSION M87924.1 GI:174871
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 107)
Sinnott, D., Richer, C., Deragon, J.-M. and Labuda, D.
Alu RNA transcripts in human embryonal carcinoma cells. Model of
post-transcriptional selection of master sequences
J. Mol. Biol. (1992) In press
Location/Qualifiers
FEATURES
source 1. 107
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Nteraz21"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"
BASE COUNT 28 a 30 c 35 g 14 t
ORIGIN

Query Match 0.3%; Score 78; DB 9; Length 107;
Best Local Similarity 85.3%; Pred. No. 0.0019;
Matches 87; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

Qy 18560 TTTTGGAGACAGCTGTGACTCTGTTGCCAGGCTGAGTGACAGTGACCATCTAGCT 18619
|||||
Db 106 TTTTGGAGACAGCTGTGACTCTGTTGCCAGGCTGAGTGACAGTGACCATCTAGCT 47
|||||

Qy 18620 CACTGACGCTCTACTCTCTGCGGTTCACAGGATTCCTCC 18661
|||||
Db 46 CACTGACGCTCTACTCTCTGCGGTTCACAGGATTCCTCC 5
|||||

RESULT 13
HUMALCE221/c 103 bp ss-RNA PRI 15-APR-1994
LOCUS HUMALCE221 Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION M87896
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 103)
Sinnott, D., Richer, C., Deragon, J.-M. and Labuda, D.
Alu RNA transcripts in human embryonal carcinoma cells. Model of
post-transcriptional selection of master sequences
J. Mol. Biol. (1992) In press
Location/Qualifiers
FEATURES
source 1. 103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="Nteraz21"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"
BASE COUNT 25 a 27 c 33 g 18 t
ORIGIN

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2000, 21:16:14 ; Search time 939.94 Seconds

(without alignments)
7719.451 Million cell updates/sec

Title: US-08-852-495C-1_COPY_168000_197000

Perfect score: 29001

Sequence: 1 TGTATTAGAGGAAAAAAGCA.....TAGATAAACCTTCTCCTT 29001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 433070

Minimum DB seq length: 10

Maximum DB seq length: 110

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N.Geneseq_36.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	72	0.2	108	X12095	Human biallelic po
2	70	0.2	108	X12095	Human biallelic po
3	67.6	0.2	108	X12095	Human gene signatu
4	65.6	0.2	100	T24892	Human gene signatu
5	63.4	0.2	88	V39744	Microsatellite ana
6	62.4	0.2	100	T24892	Human gene signatu
7	62.6	0.2	103	T26213	Human gene signatu
8	62.4	0.2	108	T25009	Human gene signatu
9	61.6	0.2	87	T21566	Human gene signatu
10	61.8	0.2	103	T20927	Human gene signatu
11	61.4	0.2	108	T26828	Human gene signatu
12	61	0.2	91	T25854	Human gene signatu
13	61	0.2	108	T26828	Human gene signatu
14	60.6	0.2	93	T22572	Human gene signatu
15	60.4	0.2	103	T26213	Human gene signatu
16	58.6	0.2	100	X12087	Human biallelic po
17	58.6	0.2	100	X12085	Human biallelic po
18	58	0.2	87	T21566	Human gene signatu
19	57.4	0.2	93	T24259	Human gene signatu
20	57.4	0.2	100	X12086	Human biallelic po
21	56.6	0.2	92	T66081	(dC-da)n. (dG-dT)n
22	56.4	0.2	93	T22572	Human gene signatu
23	55.8	0.2	93	T24259	Human gene signatu
24	55.4	0.2	81	T24093	Human gene signatu
25	55.6	0.2	95	T23131	Human gene signatu
26	55	0.2	99	T2420	Human gene signatu
27	54.6	0.2	91	T25854	Human gene signatu
28	54	0.2	60	T65762	Repeat sequence fr
29	53.8	0.2	99	T20931	Human gene signatu
30	54	0.2	100	X12087	Human biallelic po
31	54	0.2	100	X12085	Human biallelic po
32	54	0.2	100	X12086	Human biallelic po
33	53.6	0.2	69	Q29016	Probe to internal
34	53.6	0.2	91	T65740	Repeat sequence fr

C 35	53.4	0.2	99	1	T23728	Human gene signatu
C 36	52	0.2	97	1	T26728	Human gene signatu
C 37	53	0.2	69	1	Q29016	Probe to internal
C 38	51.6	0.2	103	1	T20927	Human gene signatu
C 39	51	0.2	84	1	T25848	Human gene signatu
40	51.2	0.2	93	1	T25688	Human gene signatu
41	50.8	0.2	97	1	T26728	Human gene signatu
42	50.6	0.2	110	1	T26288	Human gene signatu
C 43	50.8	0.2	110	1	T26288	Human gene signatu
C 44	50.2	0.2	64	1	Q3534	Microsatellite seq
C 45	50	0.2	75	1	T22841	Human gene signatu

ALIGNMENTS

RESULT	ID	Score	Query Match	length	ID	Description
1	X12095	72	0.2	108	X12095	Human biallelic po
2	X12095	70	0.2	108	X12095	Human biallelic po
3	X12095	67.6	0.2	108	X12095	Human gene signatu
4	X12095	65.6	0.2	100	T24892	Human gene signatu
5	X12095	63.4	0.2	88	V39744	Microsatellite ana
6	X12095	62.4	0.2	100	T24892	Human gene signatu
7	X12095	62.6	0.2	103	T26213	Human gene signatu
8	X12095	62.4	0.2	108	T25009	Human gene signatu
9	X12095	61.6	0.2	87	T21566	Human gene signatu
10	X12095	61.8	0.2	103	T20927	Human gene signatu
11	X12095	61.4	0.2	108	T26828	Human gene signatu
12	X12095	61	0.2	91	T25854	Human gene signatu
13	X12095	61	0.2	108	T26828	Human gene signatu
14	X12095	60.6	0.2	93	T22572	Human gene signatu
15	X12095	60.4	0.2	103	T26213	Human gene signatu
16	X12095	58.6	0.2	100	X12087	Human biallelic po
17	X12095	58.6	0.2	100	X12085	Human biallelic po
18	X12095	58	0.2	87	T21566	Human gene signatu
19	X12095	57.4	0.2	93	T24259	Human gene signatu
20	X12095	57.4	0.2	100	X12086	Human biallelic po
21	X12095	56.6	0.2	92	T66081	(dC-da)n. (dG-dT)n
22	X12095	56.4	0.2	93	T22572	Human gene signatu
23	X12095	55.8	0.2	93	T24259	Human gene signatu
24	X12095	55.4	0.2	81	T24093	Human gene signatu
25	X12095	55.6	0.2	95	T23131	Human gene signatu
26	X12095	55	0.2	99	T2420	Human gene signatu
27	X12095	54.6	0.2	91	T25854	Human gene signatu
28	X12095	54	0.2	60	T65762	Repeat sequence fr
29	X12095	53.8	0.2	99	T20931	Human gene signatu
30	X12095	54	0.2	100	X12087	Human biallelic po
31	X12095	54	0.2	100	X12085	Human biallelic po
32	X12095	54	0.2	100	X12086	Human biallelic po
33	X12095	53.6	0.2	69	Q29016	Probe to internal
34	X12095	53.6	0.2	91	T65740	Repeat sequence fr

Query Match 0.2%: Score 72; DB 1; Length 108;
Best local similarity 79.2%: Pred. No. 0.011;

Matches 84; Conservative 1; Mismatches 21; Indels 0; Gaps 0;

QY	8485	TTTTTTTAAATAGATGAGGATTTTGGCATTTAACGAGGAGCTTGAACCTGACCT	8544
DB	3	TTTTTTTAAATAGATGAGGATTTTGGCATTTAACGAGGAGCTTGAACCTGACCT	62
QY	8545	CAAGTATCTGCCACCTTGGCCCTTCATATGCTGGGATTTACAG	8590
DB	63	CAAGTATCTGCCACCTTGGCCCTTCATATGCTGGGATTTACAG	108

ID	Accession	Standard	DNA	Length
AC	X12095	standard	DNA	108 BP.
AC	X12095			
DE	30-MAR-1999	(first entry)		
KW	Human biallelic polymorphic DNA fragment TIGR-A003M18a.			
KW	Polymorphism; biallelic; human; forensic; paternity testing; disease;			
KW	detection; phenotypic typing; characteristic; infection; hereditary;			
KW	autoimmune disease; cancer; inflammation; drug; therapy; medication;			
KW	treatment; marker; ss.			
OS	Homo sapiens.			
PN	W09820165-A2.			
PD	14-MAY-1998.			
PE	05-NOV-1997; U20313.			
PR	06-NOV-1996; US-030455.			
PA	(MHED) WHITEHEAD INST BIOMEDICAL RES.			
PI	Hudson T, Lander ES, Wang D;			
DR	WPI; 98-286974/25.			
PT	New isolated nucleic acid segments from the human genome - used for			
PT	determining polymorphic forms for use in e.g. forensics, paternity			
PT	testing or phenotypic typing for disease			
PS	Claim 1: Page 219; 310pp; English.			
CC	X10269-X12937 are human DNA fragments which contain biallelic polymorphic			
CC	markers which have been isolated using the primers represented in			
CC	X09121-X10268. The base occupying the polymorphic site is indicated by			
CC	the appropriate IUPAC-IUB ambiguity code. These fragments can be used in			
CC	methods for determining polymorphic forms in an individual for use in			
CC	e.g. forensics, paternity testing or for phenotypic typing for diseases			
CC	such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,			
CC	muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial			
CC	hypercholesterolemia, polycystic kidney disease, hereditary			
CC	spherocytosis, von Willebrand's disease, tubercous sclerosis, hereditary			
CC	haemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos			
CC	syndrome, osteogenesis imperfecta, acute intermittent porphyria,			
CC	autoimmune diseases, inflammation, cancer, diseases of the nervous			
CC	system, infection by pathogenic microorganisms, and characteristics such			
CC	as longevity, appearance (e.g. baldness, obesity), strength, speed,			
CC	endurance, fertility, and susceptibility or receptivity to particular			
CC	drugs or therapeutic treatments. The isolated polymorphic nucleic acid			
CC	segments can also be used to produce medicaments for the treatment or			
CC	prophylaxis of such diseases.			
SQ	Sequence 108 BP; 19 A; 23 C; 28 G; 37 T;			

PI Matsubara K.; Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1748; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-126637 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 108 BP; 34 A; 31 C; 26 G; 15 T;

```

Query Match          0.2%;   Score 67.6; DB 1; Length 108;
Best Local Similarity 75.9%; Pmed. No. 0.05;
Matches      82; Conservative    0; Mismatches     26; Indels       0; Gaps        0;

OY  19044 TTTTTTTTTTTTTTTTATACGAGAGTCCTCACTGTCCATGCCAGCGTGAGTCACGC 19103
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    108 TTTGATTGTTGTTGGTTTTCACAAGGGGCTTGCCTGTGCATCATGCGGTGAATCACTG 49

OY  19104 GTGTGAACATGGCTCACATGCACGCTTCMAACTACTAGCTCAAGCATC 19151
      | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    48 GGCTGACCATGGCTCACATGCACGCTTGGGCTTCATGGGCTCAGGCGATC 1

RESULT      4
ID T24892 standard; cDNA to mRNA; 100 BP.
AC T24892;
DR 05-NOV-1996 (first entry)
DE Human gene signature HUMGS06998.
KW Human gene signature; messenger RNA; mRNA; relative abundance; frequency;
KM human; cloning; mapping; non-biased library; diagnosis; detection;
KK cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN W09514772-A1.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
DR Matsubara K, Okubo K;
PI MPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1720; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in TI9001-122637 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2000, 15:09:23 ; Search time 13753.1 Seconds
(without alignments)
8546.953 Million cell updates/sec

Title: US-08-852-495C-1_COPY_168000_197000
Perfect score: 29001
Sequence: 1 TGTTAGAGGAAAAAGCA.....TAGATAAACGTTGTCTT 29001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST: *
1: em_est1: *
2: em_est2: *
3: em_est3: *
4: em_est4: *
5: em_est5: *
6: em_est6: *
7: em_est7: *
8: em_est8: *
9: em_est9: *
10: em_est10: *
11: em_est11: *
12: em_est12: *
13: em_est13: *
14: em_est14: *
15: em_est15: *
16: em_est16: *
17: em_est17: *
18: em_est18: *
19: em_est19: *
20: gb_est1: *
21: gb_est2: *
22: gb_est3: *
23: gb_est4: *
24: gb_est5: *
25: gb_est6: *
26: gb_est7: *
27: gb_est8: *
28: gb_est9: *
29: gb_est10: *
30: gb_est11: *
31: gb_est12: *
32: gb_est13: *
33: gb_est14: *
34: gb_est15: *
35: gb_est16: *
36: gb_est17: *
37: gb_est18: *
38: gb_est19: *
39: gb_est20: *
40: gb_est21: *
41: gb_est22: *
42: gb_est23: *
43: gb_est24: *
44: gb_est25: *

45: gb_est26: *
46: gb_est27: *
47: gb_est28: *
48: gb_est29: *
49: gb_est30: *
50: gb_est31: *
51: gb_est32: *
52: em_est20: *
53: em_est21: *
54: em_est22: *
55: em_est23: *
56: em_est24: *
57: em_est25: *
58: em_est26: *
59: gb_est33: *
60: gb_est34: *
61: gb_est35: *
62: gb_est36: *
63: gb_est37: *
64: gb_est38: *
65: em_est27: *
66: em_est28: *
67: em_est29: *
68: em_est30: *
69: gb_est39: *
70: gb_est40: *
71: gb_est41: *
72: gb_est42: *
73: gb_est43: *
74: gb_est44: *
75: em_est31: *
76: em_est32: *
77: em_est33: *
78: em_est34: *
79: gb_est45: *
80: gb_est46: *
81: gb_est47: *
82: gb_est48: *
83: gb_est49: *
84: gb_est50: *
85: gb_est51: *
86: em_est35: *
87: em_est36: *
88: em_est37: *
89: em_est38: *
90: gb_est52: *
91: gb_est53: *
92: gb_est54: *
93: gb_est55: *
94: gb_est56: *
95: em_est39: *
96: em_est40: *
97: em_est41: *
98: em_est42: *
99: em_est43: *
100: em_est44: *
101: em_est45: *
102: gb_est57: *
103: gb_est58: *
104: em_est46: *
105: gb_est59: *
106: gb_est60: *
107: gb_est61: *
108: gb_est62: *
109: gb_est63: *

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of result being printed, and is derived by analysis of the total score distribution.

Result

%
Query

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	89	0.3	105	105	AQ282107	AQ282107 RPlC11-94
2	87.2	0.3	108	84	B65160	B65160 CIT-HSP-201
3	87.2	0.3	109	84	B17434	B17434 345K2.TVB C
4	87.2	0.3	109	94	AQ028426	AQ028426 CIT-HSP-2
5	86.8	0.3	106	37	AA703692	AA703692 ag81a10.r
6	86.8	0.3	109	24	N25299	N25299 yw52c09.s1
7	86.4	0.3	107	33	AA585808	AA585808 EST99495
8	85	0.3	101	35	AA583697	AA583697 nns6f10.s
9	85.2	0.3	106	63	A1991750	A1991750 wt48e01.x
10	85	0.3	109	22	H11143	H11143 ym09c06.r1
11	85	0.3	110	30	AA244245	AA244245 nc07a04.s
12	84.4	0.3	106	30	AA250812	AA250812 zs06a05.s
13	84.6	0.3	107	35	AA565533	AA565533 nk42b1.s
14	83.8	0.3	103	108	AO535244	AO535244 RPlC11-13
15	84	0.3	109	103	AQ000347	AQ000347 RPlC11-43
16	83.2	0.3	106	94	AQ046231	AQ046231 RPlC11-36
17	83.4	0.3	109	94	AQ028426	AQ028426 CIT-HSP-2
18	83.4	0.3	110	30	AA244245	AA244245 nc07a04.s
19	82.8	0.3	102	36	AA654562	AA654562 nt75f10.s
20	82.8	0.3	107	24	H67040	H67040 yu68c01.r1
21	82.8	0.3	110	106	AQ386882	AQ386882 RPlC11-13
22	82.2	0.3	103	108	AO535244	AO535244 RPlC11-13
23	82.2	0.3	106	94	AQ062963	AQ062963 CIT-HSP-2
24	82.4	0.3	109	30	AA243009	AA243009 zr25h02.s
25	81.8	0.3	101	39	AA835205	AA835205 ak64h01.s
26	82	0.3	106	105	AQ282340	AQ282340 RPlC11-80
27	81.6	0.3	104	29	AA129957	AA129957 zn68h04.r
28	81	0.3	106	44	A1749096	A1749096 qn73g09.x
29	81.2	0.3	110	106	AQ386882	AQ386882 RPlC11-13
30	80.6	0.3	103	108	AO584425	AO584425 RPlC11-14
31	80.6	0.3	104	105	AQ321855	AQ321855 RPlC11-11
32	80.4	0.3	106	38	AA812141	AA812141 ob48h02.s
33	80.4	0.3	106	106	AQ414071	AQ414071 RPlC11-11
34	80	0.3	105	30	AA218889	AA218889 zq15d04.s
35	80	0.3	107	103	AQ240182	AQ240182 CIT-HSP-2
36	79.6	0.3	102	84	B48088	B48088 RPlC11-4N6.
37	79.8	0.3	107	62	A1933497	A1933497 wmt74d02.x
38	79.8	0.3	108	84	B65160	B65160 CIT-HSP-201
39	79.8	0.3	110	39	AA897366	AA897366 am06h02.s
40	79.2	0.3	100	30	AA252633	AA252633 zq43g05.r
41	79.4	0.3	105	28	AA078003	AA078003 7H12D08 C
42	79.4	0.3	105	105	AQ276193	AQ276193 C1MB1-E1
43	79.4	0.3	107	24	H67040	H67040 yu68c01.r1
44	79.6	0.3	110	109	AQ634950	AQ634950 RPlC11-4
45	79	0.3	103	35	AA570438	AA570438 nk63g02.s

ALIGNMENTS

RESULT 1
LOCUS AQ282107 105 bp DNA GSS 27-APR-1999
DEFINITION RPlC11-94B21.r1 RPlC11-11 Homo sapiens genomic clone RPlC11-94B21,
genomic survey sequence.

ACCESSION AQ282107
VERSION AQ282107.1 GI:3907976
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 105)

AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
TITLE Use of human BAC End Sequences for Sequence-Ready Map Building
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPlC11. For BAC
library availability, please contact Pieter de Jong
(pieter@edlong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.

FEATURES
source Location/Qualifiers
1..105

/organism="Homo sapiens"
/db_xref="GB:7535756"
/db_xref="taxon:9606"
/clone="RPlC11-94B21"
/clone_11b="RPlC11-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
RPlC11 Human Male BAC library"
BASE COUNT 26 a 31 c 30 g 18 t
ORIGIN

Query Match 0.3%; Score 89; DB 105; Length 105;
Best Local Similarity 90.5%; Pred. No. 0.21;
Matches 95; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 16731 GCTCAGCCTGTATCCACGACCTTGGAGCGGACAGATCAGAGTCAGGA 16790
|||||
DB 1 GCTCAGCCTGTATCCACGACCTTGGAGCGGAGTGTGATCAGAGGCGATGA 60
QY 16791 GTTGGACCGACGACCAACATGTTGAACCCCTCTCTACTA 16835
|||||
DB 61 GTACGAGACCGACCTTACCAACATGTTGAACCCCTCTCTACTA 105

RESULT 2
LOCUS B65160 108 bp DNA GSS 21-JUN-1998
DEFINITION CIT-HSP-2017G2.TVB CIT-HSP Homo sapiens genomic clone 2017G2,
genomic survey sequence.
ACCESSION B65160
VERSION B65160.1 GI:2639138
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 108)

AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building
JOURNAL Unpublished (1997)
COMMENT other GSSs: CIT-HSP-2017G2.TVB

CONTACT: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208

Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.

FEATURES
source Location/Qualifiers
1..108

[illegible]

RESULT	3
LOCUS	B17434
DEFINITION	B17434 109 bp DNA GSS 04-JUN-1998 345K2.TVB CTR978SKA1 Homo sapiens genomic clone A-345K02, genomic survey sequence.
ACCESSION	B17434
VERSION	B17434.1 GI:2125183
KEYWORDS	GSS.
SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 109) Adams,M.D., Kelley,J.M., Rounsley,S.R. and Venter,J.C. Use of a BAC end Sequence Database for Sequence-Ready Map Building Unpublished (1997)
TITLE	Other_GSSs: 345K02.tp 345K02.TPB
JOURNAL	Contact: Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel.: 301 838 0200 Fax: 301 838 0208 Email: mdamas@tigr.org Clones are available from Research Genetics (info@resgen.com), BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html Seq primer: 17
COMMENT	Class: BAC ends.
FEATURES	Location/Qualifiers
source	1..109 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="A-345K02" /clone_1lb="CTR978SKA1" /sex="Female" /cell_type="Fibroblast" /note="Vector: pBAC108L, Site_1: HindIII, Site_2: HindIII Caltech Human BAC Library AT"
BASE COUNT	24 a 30 c 31 g 24 t
ORIGIN	

Query Match	0.38;	Score 87.2;	DB 84;	Length 109;
Best Local Similarity	88.0%;	Pred. No. 0.33;		
Matches 95;	Conservative 0;	Mismatches 13;	Indels 0;	Gaps 0;

Accession	Sequence	Position
Oy	GGCTATACCTTATCCAGACATTTGGGAACATGAGTGGTGGATCACTTATGGTCA	27476
Db	GGCTATACCTTATCCAGACATTTGGGAACATGAGTGGTGGATCACTTATGGTCA	61
Oy	GGATTTCAAAACCAAGCTTGCCCAACATGTTGAAAAACCATCTCTCTACTA	27524
Db	GGATTTCAAGACCAAGCTTGCCCAACATGTTGAAAAACCATCTCTCTACTA	109

RESULT	4				
LOCUS	AA028426				
DEFINITION	AA028426	109 bp	DNA	SS	30-JUN-1998
ACCESSION	CIT-HSP-2313G15	TF	CIT-HSP	Homo sapiens genomic clone	2313G15,
VERSION	AA028426				
KEYWORDS	AA028426.1	GI:3268648			
SOURCE	GSS.				
ORGANISM	human.				
REFERENCE	Homo sapiens				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
	1 (bases 1 to 109)				
	Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.				
TITLE	Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)				
JOURNAL	Unpublished (1998)				
COMMENT	Contact: Mark Adams The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: mdadams@tigr.org Clones are available from Research Genetics (Info@resgen.com). BAC end search page: http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html. Seq primer: M13-21 Class: BAC ends.				

FEATURES					
source	location/Qualifiers				
	1..109	/organism="Homo sapiens"			
	/db_xref="taxon:9606"				
	/clone="2313G15"				
	/clone_1lb="CIR-HSP"				
	/sex="Male"				
	/cell_type="Sperm"				
	/note="Vector: pBelobAC11; site_1: HindIII; site_2: HindIII"				
BASE COUNT	19 a	36 c	25 g	29 t	
ORIGIN					
Query Match	0.3%;	Score 87.2;	DB 94;	Length 109;	
Best Local Similarity	88.0%;	Pred. No. 0.33;			
Matches	95;	Conservative 0;	Mismatches 13;	Indels 0;	Gaps 0
OY	22365 TTTTTCGTGAATGGAGTCTCACTCTGTTTGCCCAAGCGTAGTACAGTGGCACAAATCTTG 22424				
Dd	2 TGTTTTCTGAGCAGCATCTCACTCTGTGCACCACGAGCTGGAGTGCAGTGGCACAGTCTGA 61				
OY	22425 GTTCACTGCAACTGCCAATTCTGTGGTTTAACGCACTCTCTGCATCA 22472				
Dd	62 GCTCACTGCAACTCCACCTCTGTGGTTTAACGCAATCTCTCTGCATCA 109				

RESULT	5			
AA703692/C				
LOCUS				
DEFINITION	AA703692	106 bp	mRNA	EST
	ag6ta10.r1	Stratagene	hNT neuron (#93723)	Homo sapiens cDNA clone
	IMAGE:1140858	5'	similar to contains	Alu repetitive element; mRNA sequence.

ACCESSION AA703692 GI:2713610
 VERSION AA703692.1
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 106)
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Giesel, G., Jost, S., Kitzman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Mattin, J., Moore, B., Scheinberg, K., Stepien, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R. WashU-NCI human EST Project
 JOURNAL Unpublished (1997)
 COMMENT On Sep 12, 1996 this sequence version replaced gi:1397630.
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -28ml3 rev1 ET from Amersham
 High quality sequence stop: 53.
 FEATURES
 SOURCE
 1..106
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1140858"
 /lab_host="Stratagene hNT neuron (#937233)"
 /dev_stage="hNT neurons"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Differentiated, post mitotic hNT neurons. Average insert size: 1.5 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGCGCAGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"
 BASE COUNT 19 a 29 c 29 g 29 t
 ORIGIN
 Query Match 0.3%; Score 86.8; DB 37; Length 106;
 Best Local Similarity 88.7%; Pred. No. 0.37;
 Matches 94; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
 Oy 16734 CACGCTGTATCCAGACCTTTGGAGGCGGAGGAGATCAGATCAGAGT 16793
 ||||||||||| 1 ||||||||||| 1 ||||||||||| 1 ||||||||||| 1
 Db 106 CACGCTGTATCCAGACCTTTGGAGGCGGAGGAGATCAGATCAGAGT 47
 Oy 16794 TGAGACGAGCTGACCAATGATGTAACCCGTCTCTCTAACA 16839
 ||||||| 1 |||| 1 |||| 1 |||| 1 |||| 1 |||| 1 |||| 1 |||| 1
 Db 46 CGAGACCATCTGCTGTAACGCGTAACTCTCTCTAATAA 1
 RESULT 6
 N25299 109 bp mRNA EST 28-DEC-1995
 LOCUS yw52c09.s1 Weizmann Olfactory Epithelium Homo sapiens cDNA clone
 DEFINITION IMAGE:255856 3' similar to contains Alu repetitive element; mRNA
 sequence.
 ACCESSION N25299
 VERSION N25299.1 GI:1139449
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 109)
 AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapelli, B., Chisoe, S., Dietrich, N., Dubuque, T., Favell, A., Gish, W., Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,

TITLE Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L., Rohlfing, T., Scheinberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J., Trevaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R. and Marra, M.
 JOURNAL Generation and analysis of 280,000 human expressed sequence tags
 MEDLINE Genome Res. 6 (9), 807-828 (1996)
 COMMENT On Apr 14, 1993 this sequence version replaced gi:837394.
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Source: IMAGE Consortium, LNL
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: ml3 -40 forward
 High quality sequence stop: 307.
 FEATURES
 SOURCE
 1..109
 /organism="Homo sapiens"
 /db_xref="GDB:386265"
 /db_xref="taxon:9606"
 /clone="IMAGE:255856"
 /clone_lib="Weizmann Olfactory Epithelium"
 /sex="Female"
 /tissue_type="Olfactory epithelium"
 /dev_stage="35 year Old"
 /lab_host="SOLR cells (kanamycin resistant)"
 /note="Organ: nose; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Olfactory epithelium, normal. Average insert size: 0.8 kb; Uni-ZAP XR Vector. Library constructed by N. Walker, D. Lancel, Weizmann Institute of Science. ~5' adaptor sequence: 5' GAATTCGCGCAGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"
 BASE COUNT 13 a 34 c 24 g 35 t
 ORIGIN
 Query Match 0.3%; Score 86.8; DB 24; Length 109;
 Best Local Similarity 86.2%; Pred. No. 0.37;
 Matches 94; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
 Oy 16901 AGCAGAGAAATCACTTAAACCCAGAGGAGATTCAGAGCTGAGATCGGCGAC 16960
 ||||||||||| 1 ||||||||||| 1 ||||||||||| 1 ||||||||||| 1
 Db 109 AGCAGAGAAATGCGAGAACTGGAGGAGAGATTCAGAGCTGAGATCGGCGAC 50
 Oy 16961 TGCATTCCAGCTGGAGAGAGAGAGATCCGCTCAAAATTTAA 17009
 |||| 1 |||| 1 |||| 1 |||| 1 |||| 1 |||| 1 |||| 1 |||| 1 |||| 1
 Db 49 TGCATTCCAGCTGGAGAGAGAGAGAGATCCGCTCAAAANNNAAAA 1
 RESULT 7
 AA385808 107 bp mRNA EST 21-APR-1997
 LOCUS AA385808/c
 DEFINITION EST99495 Thyroid Homo sapiens cDNA 5' end similar to EST containing
 Alu repeat, mRNA sequence.
 ACCESSION AA385808
 VERSION AA385808.1 GI:2038127
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Val, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghegan, N.S., Glodok, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,

TITLE
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence

JOURNAL
JOURNAL OF MOLECULAR EVOLUTION

PAGES
317-329

1995
1995

**JOURNAL
MEDLINE
COMMENT**

121:40200
On Jan 25, 1995 this sequence version replaced gi:637865.
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (<http://www.tigr.org/tdb/hg1/hg1.html>)
Seq primer: M3 Reverse.

FEATURES

```

/organism="Homo sapiens"
/db.xref="ATCC (host):189984"
/db.xref="taxon:9606"
/clone_lib="thyroid"
/dev_stage="adult"
/notes="Organ: thyroid gland; Vector: plasmid SK-7;
Site_1: ECoRI; Site_2: XhoI"
BASE COUNT      16 a      34 c      28 g      26 t      3 others
ORIGIN

```

Query Match	0.38	Score	86.4	DB	33	Length	107
Best Local Similarity	86.98	Pred	No. 0.41				
Matches	93	Conservative	0	Mismatches	14	Indels	0
						Gaps	0

Qy	16882	GGGACACAGGAGGACAGAGAAATCATCTTGAACCCAGGAGGACAGAGATTTGCAGTAGCTGAGA	16951
Db	107	GGGAGNCTGTGAGCGACGAGAAACGCTGAACCCAGAAAGCAGAGCTTGCAGTAGCTGAGA	48
Qy	16952	TGCGCCCACTGCATTCACACCTGGGAGACAGACGAGACTCCGCTTC	16998
Db	47	TCTGCCCACTGCATTCACACCTGGGTTCACAGGCGAGACTCTGTCTC	1

RESULT 8
AA583697/c

LOCUS	AA583697	101 bp	mRNA	EST	26-SEP-1997
DEFINITION	nm58f10.s1	NCI_CGAP_kid6	Homo sapiens	CDNA clone	IMAGE:1088107 3
	similar to contains Alu repetitive element, mRNA sequence.				

ACCESSION	AA58369/
VERSION	AA583697.1
KEYWORDS	GI:2368306
SOURCE	EST.
	human.

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia
Eutheria; Primates; Catarrhini; Hominoideae; Homo
REFERENCE
1 (bases 1 to 101)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP)
Tumor Gene Index

JOURNAL unpublished (1997)
COMMENT On Apr 14, 1993 this sequence version replaced g1:692704.

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

Email: Robert.Stratusberg@lin.gov
Tissue Procurement: D. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing
Center
Clone distribution: NCI-CGP clone distribution information
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbir/image/image.html

Insert Length: 1890 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 93.
Location/Qualifiers

FEATURES

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1088107"
/clone_id="NCI_CGAP_K1d6"
/sex="mixed"
/tissue_type="kidney tumor"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: kidney; Vector: Bluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Pooled kidney tumors. 5' adaptor sequence: 5'
GAATTGGGACACGAG 3' 3' adaptor sequence: 5'
CTCAGAGTGTGTGTGTGTGTGT 3' Average insert size: 1.0 kb."

```

Query Match	0.3%	Score 85;	DB 35;	Length 101;
Best Local Similarity	90.1%;	Pred. No. 0.61;		
Matches 91;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0;

```
Oy   16733 TCACGCCCTGTAATCCAGCACATTGGGAGGCCAAGCGGAGATCAGAGGTCAAGAGT 167922
      |||||
Db    101 TCACGCCTGTAATCCAGCACATTGGGAGGCCAAGCGGCGGATCAACAAGTCGTGAGT 42
```

```

Oy  16793  TTGAGACCAAGCCTGACCAACAATGGTGAACCCCTGTCTTAC 16833
      |||||
Db    41    TCGAGATCAGCCTGGCCAAACAATGTGAAATCCGTTTCTAC 1

```

RESULT 9
AI991750/C

LOCUS	106 bp	mRNA	EST	08-SEP-1999
DEFINITION	w48e01.x1	NCI_CGAP_Pan1 Homo sapiens CDNA IMAGE:2510712 3		
		similar to contains Alu repetitive element; contains element LTR8 repetitive element ; mRNA sequence.		

ACCESSION	AI991750
VERSION	AI991750.1
	GI:5838578

SOURCE ORGANISM

REFERENCE
1 (bases 1 to 106)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia
Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS NCI-CGAP <http://www.ncd.nih.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
JOURNAL Tumor Gene Index
Unpublished (1997)

COMMENT
On Dec 20, 1995 this sequence version replaced g1:1133359.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nh.gov

Life technologies catalog #: 11548-013
DNA sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/Image/Image.html

Seq primer: -40UP from Glibco

High quality sequence stop: 62.
Location/Quality

FEATURES	SOURCE
location/Qualifiers	1. .106
/organism="Homo sapiens"	/db_xref="taxon:9606"
/clone="IMAGE:2510712"	/clone_lib="NCI_CGAP_Pa1"
/tissue_type="adenocarcinoma"	/lab_host="DH10B"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site_1: SalI Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.72 kb. Life Technologies catalog #:	11548-013"
BASE COUNT	24 a 23 c 22 g 37 t
ORIGIN	

```

/sex="female"
/dev_stage="73 days post natal"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: whole brain; Vector: latmid BA; Site_1: Not
I; Site_2: Hind III; 1st strand cDNA was primed with a Not
I - oligo(dT) primer [5',
ACTGTGAAGATTCGCGCCGACAGGAATTTTTTTTTTTTTTTT 3'];
double-stranded cDNA was ligated to Hind III adaptors
(Pharmacia), digested with Not I and directionally cloned
into the Not I and Hind III sites of the latmid BA vector.
Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."

```

	Query Match	0.3%	Score 85;	DB 22;	Length 109;
	Best Local Similarity	86.2%	Pred. No. 0.59;		
	Matches	94;	Conservative	0;	Mismatches 15; Indels 0; Gaps 0;
OY	310	TGAGCGGAGNCTCACTCTGTCACCCAGGCTGGAGTGCAGTACGCTATCCGGGCTCACT	369		
Db	109	TGAGAGGCGNCTCACTCTGTATCCAGGCTGGAGTGCAGTACGCTGATCTACGCTCACT	50		
OY	370	GCACCTCCGCTCCCGGGTTCAGGCAATTCCTGCTCAGCTCCCGCA	418		
Db	49	GGAACCTCCGCTCCCGGGTTCAGGCAATTCCTCAGCTCCCGTGA	1		

```
/organism="Homo sapiens"  
/db_xref="GDB:419851"  
/db_xref="taxon:9606"  
/clone="IMAGE:47310"  
/clone_lib="Soares infant brain 1N1B"
```

```

1. 106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:564368"
/clone_lib="NCI_GCAP-CGB1"
/tissue_type="NCI_GCAP-CGB1"
/lisuse_type="germinal center B cell"
/lisuse_DHI08
/Note=Vector: pTV73D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand
cDNA was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IgG-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Mark (GBER). cDNA synthesis was
primed with a Not I - oligo(dG) primer
15'-GTGACCAATCAAGATCGAGAGCGCCGCTATTTTTTTTTTTT-
3']. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pTV73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

```

Query Match	0.38;	Score 84.6;	DB 35;	Length 107;
Best Local Similarity	86.98;	Pred. NO. 0.66;		
Matches 93;	Conservative 0;	Mismatches 14;	Indels 0;	Gaps 0;
By 16864 TGGCATCTGCGCTGTAGTCCCACTACTCGGACACTGAGAGGACAGGACAAATCACTTGAAACC 16923				

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 16, 2000, 20:15:34 ; Search time 593.69 Seconds
(without alignments)
6349.604 Million cell updates/sec

Title: US-08-852-495C-1_COPY_168000_197000
Perfect score: 29001
Sequence: 1 TGTTCAGAGCAAAAAGCA.....TAGATAAACGTTGTCCTT 29001

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/prodata/1/lna/5A_COMB.seq:*
2: /cgn2_6/prodata/1/lna/5B_COMB.seq:*
3: /cgn2_6/prodata/1/lna/5C_COMB.seq:*
4: /cgn2_6/prodata/1/lna/5D_COMB.seq:*
5: /cgn2_6/prodata/1/lna/6_COMB.seq:*
6: /cgn2_6/prodata/1/lna/PCrUS_COMB.seq:*
7: /cgn2_6/prodata/1/lna/Backfillseq1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	82.6	0.3	105	4	US-08-481-658B-65
C 2	82.6	0.3	105	4	US-08-477-504A-65
C 3	82.6	0.3	105	4	US-08-486-756A-65
C 4	82.6	0.3	105	4	US-08-485-862B-65
C 5	82.6	0.3	105	5	US-08-787-739-65
C 6	68.6	0.2	84	3	US-08-454-557C-91
C 7	68.6	0.2	84	4	US-08-340-426D-91
C 8	68.6	0.2	84	4	US-08-450-673C-91
C 9	68.6	0.2	84	6	PCT-US95-17111A-91
C 10	64.8	0.2	105	4	US-08-481-658B-65
C 11	64.8	0.2	105	4	US-08-477-504A-65
C 12	64.8	0.2	105	4	US-08-486-756A-65
C 13	64.8	0.2	105	4	US-08-485-862B-65
C 14	59.4	0.2	105	5	US-08-787-739-65
C 15	59.4	0.2	78	3	US-08-454-557C-70
C 16	59.4	0.2	78	4	US-08-340-426D-70
C 17	59.4	0.2	78	4	US-08-450-673C-70
C 18	59.4	0.2	78	6	PCT-US95-17111A-70
C 19	57.8	0.2	78	3	US-08-454-557C-70
C 20	57.8	0.2	78	4	US-08-340-426D-70
C 21	57.8	0.2	78	6	PCT-US95-17111A-70
C 22	57.8	0.2	78	6	PCT-US95-17111A-70
C 23	56.6	0.2	92	1	US-08-222-177A-930
C 24	56.6	0.2	85	3	US-08-454-557C-92
C 25	56	0.2	85	4	US-08-340-426D-92
C 26	56	0.2	85	4	US-08-450-673C-92
C 27	56	0.2	85	6	PCT-US95-17111A-92

28	54	0.2	60	1	US-08-222-177A-244	Sequence 244, App
29	53.6	0.2	91	1	US-08-222-177A-166	Sequence 166, App
30	52.4	0.2	84	3	US-08-454-557C-91	Sequence 91, App1
31	52.4	0.2	84	4	US-08-340-426D-91	Sequence 91, App1
32	52.4	0.2	84	4	US-08-450-673C-91	Sequence 91, App1
33	52.4	0.2	84	6	PCT-US95-17111A-91	Sequence 91, App1
34	50.4	0.2	60	3	US-08-454-557C-60	Sequence 60, App1
35	50.4	0.2	60	4	US-08-340-426D-60	Sequence 60, App1
36	50.4	0.2	60	4	US-08-450-673C-60	Sequence 60, App1
37	50.4	0.2	60	6	PCT-US95-17111A-60	Sequence 60, App1
38	50	0.2	76	3	US-08-454-557C-69	Sequence 69, App1
39	50	0.2	76	4	US-08-340-426D-69	Sequence 69, App1
40	50	0.2	76	4	US-08-450-673C-69	Sequence 69, App1
41	50	0.2	76	6	PCT-US95-17111A-69	Sequence 69, App1
42	50	0.2	83	4	US-08-481-658B-66	Sequence 66, App1
43	50	0.2	83	4	US-08-477-504A-66	Sequence 66, App1
44	50	0.2	83	4	US-08-486-756A-66	Sequence 66, App1
45	50	0.2	83	4	US-08-485-862B-66	Sequence 66, App1

ALIGNMENTS

RESULT 1
US-08-481-658B-65/c
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OR INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/481,658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-481-658B-65
Query Match 0.3%; Score 82.6; DB 4; Length 105;
Best Local Similarity 86.7%; Pred. No. 1.7e-09;

Matches 91: Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 16744 ATCCGACACTTTGGAGAGCCGAGGAGATCAGAGTTCAGAGTTTGAGACGAC 16803

DB 105 ATCCGACACTTTGGAGAGCCGAGGAGTGTGATCACAAGGTCAGAGTTTGAGACGAC 46

OY 16804 CTGACCAACATGTGTAACCCCTGTCTCTACTAACAATAATACAAA 16848

DB 45 CTGGCCAATATGTGTGAACCCCTGTCTCTACTAACAATGTAAAAA 1

RESULT 2

US-08-477-504A-65/c

; Sequence 65, Application US/08477504A

; Patent No. 5972353

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/477,504A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3D

; TELECOMMUNICATION INFORMATION:

; TELEFAX: 415-435-2034

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-477-504A-65

Query Match 0.3%; Score 82.6; DB 4; Length 105;

Best Local Similarity 86.7%; Pred. No. 1.7e-09;

Matches 91: Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 16744 ATCCGACACTTTGGAGAGCCGAGGAGATCAGAGTTCAGAGTTTGAGACGAC 16803

DB 105 ATCCGACACTTTGGAGAGCCGAGGAGTGTGATCACAAGGTCAGAGTTTGAGACGAC 46

OY 16804 CTGACCAACATGTGTAACCCCTGTCTCTACTAACAATAATACAAA 16848

DB 45 CTGGCCAATATGTGTGAACCCCTGTCTCTACTAACAATGTAAAAA 1

RESULT 3

US-08-486-756A-65/c

; Sequence 65, Application US/08486756A

; Patent No. 5981711

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,756A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3C

; TELECOMMUNICATION INFORMATION:

; TELEFAX: 415-435-2034

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

US-08-486-756A-65

Query Match 0.3%; Score 82.6; DB 4; Length 105;

Best Local Similarity 86.7%; Pred. No. 1.7e-09;

Matches 91: Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 16744 ATCCGACACTTTGGAGAGCCGAGGAGATCAGAGTTCAGAGTTTGAGACGAC 16803

DB 105 ATCCGACACTTTGGAGAGCCGAGGAGTGTGATCACAAGGTCAGAGTTTGAGACGAC 46

OY 16804 CTGACCAACATGTGTAACCCCTGTCTCTACTAACAATAATACAAA 16848

DB 45 CTGGCCAATATGTGTGAACCCCTGTCTCTACTAACAATGTAAAAA 1

RESULT 4

US-08-485-862B-65/c

; Sequence 65, Application US/08485862B

; Patent No. 5989838

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court


```

: CITY: Tiburon
: STATE: California
: COUNTRY: USA
: ZIP: 94920
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/485,862B
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: APPLICATION DATA:
: APPLICATION NUMBER: US 08/477,504
: FILING DATE: 07-JUN-1995
: APPLICATION NUMBER: US 08/260,190
: FILING DATE: 15-JUN-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Lauder, Leona L.
: REGISTRATION NUMBER: 30,863
: REFERENCE/DOCKET NUMBER: D-0021.3D
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-435-2034
: TELEFAX: 415-435-0727
: INFORMATION FOR SEQ ID NO: 65:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 105 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOHETICAL: NO
: ANTI-SENSE: NO
:
: US-08-485-862B-65
:
Query Match      0.3%; Score 82.6; DB 4; Length 105;
Best Local Similarity 86.7%; Pred. No. 1.7e-09;
Matches 91; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
:
QY 16744 ATCCGACGACTTTGGGAGCGCGGACGATCAGAGTGTGAGACCAGC 16803
: |||||
Db 105 ATCCGACGACTTTGGGAGCGCGGAGCGTGTGATCACAAGTGTGAGAGCAGC 46
: |||||
QY 16804 CTGACCAACATGGTGAACCCCTGCTCTACTAACAATAATCAAAA 16848
: |||||
Db 45 CTGCGCAATATGTGTGAACCCCTGCTCTACTAAGAATGTAAAAA 1
:
RESULT 5
US-08-787-739-65/c
: Sequence 65, Application US/08/87739
: Patent No. 6027887
: GENERAL INFORMATION:
: APPLICANT: Zavada, Jan
: APPLICANT: Pastorekova, Silvia
: APPLICANT: Pastorek, Jaromir
: TITLE OF INVENTION: MN Gene and Protein
: NUMBER OF SEQUENCES: 96
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Leona L. Lauder
: STREET: 369 Pine Street, Suite 610
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
: CURRENT APPLICATION DATA:

```

```

: APPLICATION NUMBER: US/08/787,739
: FILING DATE: 24-JAN-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/485,049
: FILING DATE: 07-JUN-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/486,756
: FILING DATE: 07-JUN-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/477,504
: FILING DATE: 07-JUN-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/481,658
: FILING DATE: 07-JUN-1995
: APPLICATION DATA:
: APPLICATION NUMBER: US 08/485,862
: FILING DATE: 07-JUN-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/485,863
: FILING DATE: 07-JUN-1995
: APPLICATION DATA:
: APPLICATION NUMBER: US 08/487,077
: FILING DATE: 07-JUN-1995
: ATTORNEY/AGENT INFORMATION:
: NAME: Lauder, Leona L.
: REGISTRATION NUMBER: 30,863
: REFERENCE/DOCKET NUMBER: D-0021.4
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-981-2034
: TELEFAX: 415-981-0332
: INFORMATION FOR SEQ ID NO: 65:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 105 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOHETICAL: NO
: ANTI-SENSE: NO
:
: US-08-787-739-65
:
Query Match      0.3%; Score 82.6; DB 5; Length 105;
Best Local Similarity 86.7%; Pred. No. 1.7e-09;
Matches 91; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
:
QY 16744 ATCCGACGACTTTGGGAGCGCGGACGATCAGAGTGTGAGACCAGC 16803
: |||||
Db 105 ATCCGACGACTTTGGGAGCGCGGAGCGTGTGATCACAAGTGTGAGAGCAGC 46
: |||||
QY 16804 CTGACCAACATGGTGAACCCCTGCTCTACTAACAATAATCAAAA 16848
: |||||
Db 45 CTGCGCAATATGTGTGAACCCCTGCTCTACTAAGAATGTAAAAA 1
:
RESULT 6
US-08-454-557C-91/c
: Sequence 91, Application US/08/454557C
: Patent No. 5830670
: GENERAL INFORMATION:
: APPLICANT: de la Monte, Suzanne
: APPLICANT: Wands, Jack R.
: TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
: OF ALZHEIMER'S DISEASE
: NUMBER OF SEQUENCES: 121
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
: STREET: 1100 New York Avenue, Suite 600
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20005-3934
: COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-91

Query Match 0.2%; Score 68.6; DB 3; Length 84;
Best Local Similarity 89.2%; Pred. No. 1.8e-06;
Matches 74; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 16734 CACGCTGTATCCCGACACTTTGGAGGCCAAGCGGACAGATCCAGAGTTCAGAGTT 16793
DB 83 CACGCTGTATCCCGACACTTTGGAGGCCGCGGCGGATCCAGAGTTCAGAGTT 24

QY 16794 TGAGACACAGCTGACCAATGG 16816
DB 23 CGACACACAGCTGATGACATGG 1

RESULT 7
US-08-340-426D-91/c
Sequence 91, Application US/08340426D
Patent No. 5948634
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,426D
FILING DATE: 14-NOV-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 base pairs

TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-340-426D-91

Query Match 0.2%; Score 68.6; DB 4; Length 84;
Best Local Similarity 89.2%; Pred. No. 1.8e-06;
Matches 74; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 16734 CACGCTGTATCCCGACACTTTGGAGGCCAAGCGGACAGATCCAGAGTTCAGAGTT 16793
DB 83 CACGCTGTATCCCGACACTTTGGAGGCCGCGGCGGATCCAGAGTTCAGAGTT 24

QY 16794 TGAGACACAGCTGACCAATGG 16816
DB 23 CGACACACAGCTGATGACATGG 1

RESULT 8
US-08-450-673C-91/c
Sequence 91, Application US/08450673C
Patent No. 5948888
GENERAL INFORMATION:
APPLICANT: de la Monte, Suzanne
TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
TITLE OF INVENTION: of Alzheimer's Disease
NUMBER OF SEQUENCES: 121
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Avenue, Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/450,673C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-450-673C-91

Query Match 0.2%; Score 68.6; DB 4; Length 84;
Best Local Similarity 89.2%; Pred. No. 1.8e-06;
Matches 74; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 16734 CACGCTGTATCCCGACACTTTGGAGGCCAAGCGGACAGATCCAGAGTTCAGAGTT 16793
DB 83 CACGCTGTATCCCGACACTTTGGAGGCCGCGGCGGATCCAGAGTTCAGAGTT 24

QY 16794 TGAGACACAGCTGACCAATGG 16816
DB 23 CGACACACAGCTGATGACATGG 1

Best Local Similarity 82.1%; Pred. No. 1.4e-05;
Matches 87; Conservative 0; Mismatches 17; Indels 2; Gaps 1;

QY 15096 TTTTGTATTATTAGTAGAGATGGGGTTTGGCCATGTTGGCCAGCGTGTCTGCACTCC 15155
||||| ||| ||||||||| ||||||| ||| |||||||
Db 2 TTTTACATCTTTAGTAGAGACAGAGGTTTCAACCATATTGGCCAGCGTGTCTCAAACTCC 61

OY 15156 TGGCCTCAGAGCATCTGCTGGCTGCCCTCCCAAGTGTAGGAT 15201
|| ||| ||||||| ||| ||||||||| ||||| |||||
Db 62 TGACCT--TGATGCACACAGCTCGGCTCCCAAGTGTGGGAT 105

RESULT 14

US-08-787-739-65
; Sequence 65; Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvyia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; City: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO

ANTI-SENSE: NO
US-08-787-739-65

Query Match 0.2%; Score 64.8; DB 5; Length 105;
Best Local Similarity 82.1%; Pred. No. 1.4e-05;
Matches 87; Conservative 0; Mismatches 17; Indels 2; Gaps 1;

QY 15096 TTTTGTATTATTAGTAGAGATGGGGTTTGGCCATGTTGGCCAGCGTGTCTGCACTCC 15155
||||| ||| ||||||||| ||||||| ||| |||||||
Db 2 TTTTACATCTTTAGTAGAGACAGAGGTTTCAACCATATTGGCCAGCGTGTCTCAAACTCC 61

OY 15156 TGGCCTCAGAGCATCTGCTGGCTGCCCTCCCAAGTGTAGGAT 15201
|| ||| ||||||| ||| ||||||||| ||||| |||||
Db 62 TGACCT--TGATGCACACAGCTCGGCTCCCAAGTGTGGGAT 105

RESULT 15

US-08-454-557C-70
; Sequence 70; Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; City: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-454-557C-70

Query Match 0.2%; Score 59.4; DB 3; Length 78;
Best Local Similarity 85.7%; Pred. No. 0.00018;
Matches 66; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 15081 ACCATGCTGGCTAATTTTGTATTATTAGTAGAGATGGGGTTTGGCCATGTTGGCCAGG 15140
||||| ||| ||||||||| ||||||| ||| |||||||
Db 1 ACAAGCCACAGCTAATTTTGTATTATTAGTAGAGATGGGGTTTGGCCATGTTGGCCAGG 60

OY 15141 CTGGTCTGCAACTCTG 15157
||||| |||||||||
Db 61 CTGGTCTGCAACTCTG 77

Wed Jun 21 14:43:31 2000

us-08-852-495c-1_copy_168000_197000.rni

Page 8

Search completed: June 17, 2000, 10:50:11
Job time: 251741 sec

• • •

GenCore version 4.5
Copyright (c) 1993 - 2000 Compen Ltd.

OM nucleic - nucleic search, using sv model

Run on: June 17, 2000, 10:40:12 ; Search time 29137.4 Seconds

(Without alignments)
-600.987 Million cell updates/sec

Title: US-08-852-495c-1_COPY_196000_214000
Perfect score: 18001
Sequence: 1 GATAGGCTCACTTACCA.....CCATCAGCTTCTACCTCT 18001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 882769 seqs, -486395729 residues

Total number of hits satisfying chosen parameters: 370290

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl1:*
8: gb_pl2:*
9: gb_pl3:*
10: gb_pr1:*
11: gb_pr2:*
12: gb_pr3:*
13: gb_ro:*
14: gb_sy:*
15: gb_un:*
16: gb_vl:*
17: em_fun:*
18: em_hun1:*
19: em_hun2:*
20: em_in:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_sy:*
30: em_un:*
31: em_vl:*
32: gb_htg1:*
33: gb_htg2:*
34: gb_in1:*
35: gb_in2:*
36: em_ba1:*
37: em_ba2:*
38: em_hun3:*
39: em_hun4:*
40: gb_pr4:*
41: gb_htg3:*
42: gb_htg4:*
43: gb_htg5:*
44: gb_htg6:*

45: gb_htg7:*
46: em_htg1:*
47: em_htg2:*
48: em_htg3:*
49: em_hun5:*
50: gb_pl3:*
51: gb_pl5:*
52: gb_htg8:*
53: gb_htg9:*
54: gb_htg10:*
55: gb_htg11:*
56: gb_htg12:*
57: gb_htg13:*
58: gb_htg14:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	88.8	0.5	108	10	HSUCLR2
C 2	87.4	0.5	108	11	HSUCLR2
C 3	84.4	0.5	103	9	HUMALCE221
C 4	83	0.5	108	10	HSUCLR12
C 5	82.2	0.5	107	9	HUMALCE162
C 6	79.8	0.4	108	10	HSUCLR1
C 7	79.8	0.4	108	10	HSUCLR2
C 8	79.2	0.4	108	10	HSUCLR2
C 9	77.8	0.4	108	11	HSUCLR2
C 10	77.2	0.4	91	13	HUMOTR164A
C 11	77.2	0.4	103	13	HSUCLR
C 12	76.2	0.4	108	11	HSUCLR808
C 13	75	0.4	103	13	HSUCLR
C 14	74.4	0.4	110	9	HUMALCE43
C 15	74.2	0.4	108	9	HUMALCE3M5
C 16	74	0.4	103	9	HUMALCE221
C 17	74	0.4	110	11	HSUCLR807
C 18	73.2	0.4	104	9	HUMALCE272
C 19	72	0.4	106	13	G32743
C 20	72	0.4	108	13	G43535
C 21	71.2	0.4	97	9	HUMALCE272
C 22	70.4	0.4	108	10	HSUCLR2
C 23	70.2	0.4	108	10	HSUCLR1
C 24	70.2	0.4	108	10	HSUCLR2
C 25	70.2	0.4	110	11	HSUCLR807
C 26	69.8	0.4	90	9	HUMALCE272
C 27	69.8	0.4	104	9	HUMALCE272
C 28	69.4	0.4	101	10	S79560
C 29	69.2	0.4	97	9	HUMALCE272
C 30	68.6	0.4	108	11	HSUCLR803
C 31	68.6	0.4	100	13	HUMALCE272
C 32	68.2	0.4	107	11	HSUCLR806
C 33	68	0.4	108	13	G43535
C 34	67.8	0.4	108	9	HUMALCE272
C 35	67.2	0.4	80	9	HUMALCE272
C 36	67.2	0.4	97	9	HUMALCE272
C 37	67.2	0.4	99	13	HUMALCE272
C 38	67	0.4	97	9	HUMALCE272
C 39	66.8	0.4	94	9	HUMALCE272
C 40	66.8	0.4	100	11	HSUCLR806
C 41	66.6	0.4	101	10	S79560
C 42	66.6	0.4	108	11	HSUCLR808
C 43	66	0.4	94	9	HUMALCE272
C 44	65.2	0.4	95	13	HUMALCE272
C 45	65	0.4	98	13	G33095

ALIGNMENTS

RESULT	1	HSIDLIN2	108 bp	DNA	PRI	20-MAY-1992
LOCUS		Human LDL-receptor gene	Intron 14	fragment (normal gene).		
ACCESSION		X05250				
VERSION		X05250.1	GI:34337			
KEYWORDS		Alu repetitive sequence; low density lipoprotein receptor.				
SOURCE		human.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;				
AUTHORS		Primates; Catarrhini; Homiidae; Homo.				
		1 (bases 1 to 108)				
TITLE		Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,				
		Williamson,R., and Humphries,S.				
		Unequal crossing-over between two alu-repetitive DNA sequences in				
		the low-density-lipoprotein-receptor gene. A possible mechanism for				
		the defect in a patient with familial hypercholesterolaemia				
		Eur. J. Biochem. 164 (1), 77-81 (1987)				
JOURNAL		87161901				
MEDLINE		See X05252 for deletion junction				
COMMENT		Data kindly reviewed (07-DEC-1987) by HOMPHRIS S.				
FEATURES		Location/Qualifiers				
SOURCE		1..108				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
Intron		1..108				
		/note="Intron XIV fragment"				
BASE COUNT		28 a 23 c 39 g 18 t				
ORIGIN						
Query Match		0.5%; Score 88.8; DB 10; Length 108;				
Best Local Similarity		88.9%; Pred. No. 0.00015;				
Matches		96; Conservative 0; Mismatches 12; Indels 0; Gaps 0;				
OY	13470	CTTGCGTCATTGCACACCTCGCTCGCTGGGTCACAGTAGTTCCTGACTGACCTCCCG	13529			
Db	108	CTCGGCTCAGTGCACACCTCGCTCGCTGGGTCACAGCAATTCCTCGCTCAGCCTCCG	49			
OY	13530	AGTAGCTGGGATTTCACAGCATCATCACCATCCGCGGAAATTTTGT	13577			
Db	48	ACTACTGGGATTACAGCAGCAGCTGCACACGCGCTGCGTAATTTTGT	1			
RESULT	2					
LOCUS		HSU67803	108 bp	RNA	PRI	01-AUG-1997
DEFINITION		Human small cytoplasmic Alu transcript.				
ACCESSION		U67803				
VERSION		U67803.1	GI:2289917			
KEYWORDS		Alu.				
SOURCE		human.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;				
AUTHORS		Eutheria; Primates; Catarrhini; Homiidae; Homo.				
TITLE		1 (bases 1 to 108)				
		Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.				
		cDNAs derived from primary and small cytoplasmic Alu (scAlu)				
		transcripts				
		J. Mol. Biol. 271 (2), 222-234 (1997)				
JOURNAL		97415756				
MEDLINE		2 (bases 1 to 108)				
REFERENCE		Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.				
AUTHORS		Direct Submission				
TITLE		Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The				
JOURNAL		Children's Hospital of Philadelphia, 1004F Abramson Research				
		Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA				
FEATURES		Location/Qualifiers				
SOURCE		1..108				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
		/clone="TscAlu2"				

[illegible]

Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke, B., Beisiegel, U., Dunning, A., Haviga, J. R.,
Williamson, R. and Humphries, S.

TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia

JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901

COMMENT see X05249 for deletion junction
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.

FEATURES Location/Qualifiers
1..108

source /organism="Homo sapiens"
/db_xref="taxon:9606"

misc_feature complement(1..65)
/note="Alu repeat"

Intron 1..108
/note="Intron XII fragment"

BASE COUNT 21 a 38 c 20 g 29 t
ORIGIN

Query Match 0.5%; Score 83; DB 10; Length 108;
Best Local Similarity 86.0%; Pred. No. 0.0011;

Matches 92; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 11214 TTGGCTCACTGCACCTTGGCTTGGGTTCAAGCAATATCTCCAGCCTCTTA 11273
1 11111 11111 11111 11111 11111 11111 11111 11111 11111
Db 2 TCGCCACACACACACTGCTGCTCGGTTCAACCATTTCTGCTCAGCCTCTTA 61

QY 11274 GTAGCAGGCTTACAGCATGTGCCACACACCCGGCTGTTTGTGA 11320
11111 11111 11111 11111 11111 11111 11111 11111 11111
Db 62 GTAGCTGGGATTACAGCATGTGCCACACCCGGCTGTTTGTGA 108

RESULT 5

LOCUS HUMALCE162 107 bp ss-RNA PRI 15-APR-1994

DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE162.
ACCESSION M87924.1 GI:174871

KEYWORDS Alu repeat.

SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 107)
AUTHORS Eutheria; Primates; Catarrhini; Homiidae; Homo.

TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia

JOURNAL Eur. J. Mol. Biol. (1992) In Press
FEATURES Location/Qualifiers
1..107

source /organism="Homo sapiens"
/db_xref="taxon:9606"

misc_feature complement(1..65)
/note="Alu repeat"

Intron 1..108
/note="Intron XII fragment"

BASE COUNT 28 a 30 c 35 g 14 t
ORIGIN

Query Match 0.5%; Score 82.2; DB 9; Length 107;
Best Local Similarity 87.4%; Pred. No. 0.0015;

Matches 90; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 12712 TTTTGTGAGTCTGCTGCTGCTCAAGCTGAGTGCAGTGTGCTGCGC 12771
11111 11111 11111 11111 11111 11111 11111 11111 11111
Db 107 TTTTGTGAGTCTGCTGCTGCTGCTGCGCAGTGTGAGTGCAGTGTGCGC 48

QY 12772 TCACGTCAACCTCTGCTCCGGTTCAAGTATCTCCGCC 12814
11111111 1111111111111111 1111111111
Db 47 TCACGTCAACCTCTGCTCCGGTTCAAGTATCTCTTCCGCC 5

RESULT 6

LOCUS HSLDLR1 108 bp DNA PRI 20-MAY-1992

DEFINITION Human LDL-receptor mutated gene with intron 12 deletion junction.
ACCESSION X05249

VERSION X05249.1 GI:34335

KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
human.

SOURCE

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke, B., Beisiegel, U., Dunning, A., Haviga, J. R.,
Williamson, R. and Humphries, S.

TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia

JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)
MEDLINE 87161901

COMMENT *source: hypercholesterol aemia
See X05248 for corresponding normal gene sequence
in the defective LDL-receptor gene the deletion occurred between two
alu-repetitive sequences, that are in the same direction, the
deletion eliminates exons 13 and 14 and changes the reading frame
of the resulting spliced mRNA.

Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
Location/Qualifiers
1..108

source /organism="Homo sapiens"
/db_xref="taxon:9606"

misc_feature /cell_type="blood leukocytes from a patient with familial"
1..108
/note="deletion junction region intron 12/ intron 15"

BASE COUNT 20 a 40 c 20 g 28 t
ORIGIN

Query Match 0.4%; Score 79.8; DB 10; Length 108;
Best Local Similarity 84.1%; Pred. No. 0.0034;

Matches 90; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 13471 TTGGCTCATTTGACACCTCTGCTGCTGCTTCAAGTATCTCCAGCCTCCCGA 13530
1 11111 11111 11111 11111 11111 11111 11111 11111 11111
Db 2 TCGCCACACACACACTGCTGCTCGGTTCAACCATTTCTGCTCAGCCTCCCGA 61

QY 13531 GTAGCTGGGATTACAGCATGTGCCACACCCGGCTGTTTGTGA 13577
11111 11111 11111 11111 11111 11111 11111 11111 11111
Db 62 GTAGCTGGGATTACAGCATGTGCCACACCCGGCTGTTTGTGA 108

RESULT 7

LOCUS HSLDLR2 108 bp DNA PRI 20-MAY-1992

DEFINITION Human LDL-receptor mutated gene with intron 14 deletion junction.
ACCESSION X05251

VERSION X05251.1 GI:34336

KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
human.

SOURCE

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke, B., Beisiegel, U., Dunning, A., Haviga, J. R.,
Williamson, R. and Humphries, S.

TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia

JOURNAL Eur. J. Biochem. 164 (1), 77-81 (1987)

FEATURES	source	Location/Qualifiers
COMMENT		*source: hypercholesterolemia See X05250 for corresponding normal gene sequence In the defective LDL-receptor gene the deletion occurred between two alu-repetitive sequences, that are in the same direction, the deletion eliminates exons 13 and 14 and changes the reading frame of the resulting spliced mRNA. Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
BASE COUNT	28 a 20 c 40 g 20 t	
ORIGIN		
Query Match	0.4%; Score 79.2; DB 10; Length 108;	
Best Local Similarity	83.3%; Pred. No. 0.0034;	
Matches	90; Conservative 0; Mismatches 17; Indels 0; Gaps 0;	
Db	13531	GTAGCTGGGATTACAGGCATCATCCATGCTGGATTTTGT 13577
Db	47	GTAGCTGGGATTACAGGCATCCACACGCTGCTTAATTTTGT 1
RESULT	8	
LOCUS	HSIDLRLN2	108 bp DNA PRI 20-MAY-1992
DEFINITION	Human LDL-receptor gene intron 14 fragment (normal gene).	
ACCESSION	X05250	
VERSION	X05250.1	GI:34337
KEYWORDS	Alu repetitive sequence; low density lipoprotein receptor. human.	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 108) Horsthemke,B., Beisiegel,U., Dunning,A., Haveling,J.R., Williamson,R. and Humphries,S. Unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia Eur. J. Biochem. 164 (1), 77-81 (1987)	
REFERENCE	See X05252 for deletion junction Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.	
AUTHORS	Location/Qualifiers	
TITLE	1. 108 /organism="Homo sapiens" /db_xref="taxon:9606" 1. 108 /note="intron XIV fragment"	
JOURNAL		
MEDLINE		
COMMENT		
FEATURES		
source		
intron		
BASE COUNT	28 a 23 c 39 g 18 t	
ORIGIN		
Query Match	0.4%; Score 79.2; DB 10; Length 108;	
Best Local Similarity	83.3%; Pred. No. 0.0042;	
Matches	90; Conservative 0; Mismatches 18; Indels 0; Gaps 0;	
Db	4290	ACAAAAAGTAGCAGCGCGGTGTGTCGCCACTGTATCCACGCTACTAGAGGCTGAA 4349
Db	1	ACAAAAATATAGCAGCGCGGTGTGTCGCCACTGTATCCACGCTACTAGAGGCTGAG 60
Db	4350	GGGGAGATCACTTGACCTGGAGGAGAGATTGCAATGAGCTGAG 4397

DB	61	GCAGAGAAATTGCTTGAACCCAGAGAGCGACGCTTGCAAGTGAAGCCGAG	108
RESULT	9		
LOCUS	HSU67804/C	108 bp	RNA
DEFINITION	Human small cytoplasmic Alu transcript.		PRI
ACCESSION	U67804		
VERSION	U67804.1	GI:2289918	
KEYWORDS	Alu.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 108)		
TITLE	Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L. cDNAs derived from primary and small cytoplasmic Alu (scalu) transcripts		
JOURNAL	J. Mol. Biol. 271 (2), 222-234 (1997)		
MEDLINE	97415756		
REFERENCE	2 (bases 1 to 108)		
AUTHORS	Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.		
JOURNAL	Direct Submission		
TITLE	Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The Children's Hospital of Philadelphia, 1004F Abramson Research Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA		
FEATURES	Location/Qualifiers		
source	1..108		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="TSCALU3"		
repeat_region	1..108		
	/note="scalu"		
	/rpt_family="Alu"		
	/rpt_type-dispersed		
BASE COUNT	26 a 38 c 26 g 18 t		
ORIGIN			
Query Match	0.4%; Score 77.8; DB 11; Length 108;		
Best Local Similarity	87.6%; Pred. No. 0.0069;		
Matches	85; conservative 0; Mismatches 12; Indels 0; Gaps 0;		
QY	12882	GTAGAGATGGAGTTTCCCGCTGTAGCCAGAGATGCTCTGATCTCTGACCTCGGATCC	12941
DB	97	GGAAGACGCGGTTTCACCATGTATGACCGCAGGATGCTGATCTCTGACCTTGATCC	38
OY	12942	ACCGGCTCGGCTCCCAAGTGGTGGATACAGGC	12978
DB	37	TCCCGCTTGGCCTTCACAAAGTCGTGGATACAGGC	1
RESULT	10		
LOCUS	HUMUT8164/C		
DEFINITION	Human STS UT8164, 5' primer bind, sequence tagged site.		
ACCESSION	L30244		
VERSION	L30244.1	GI:605447	
KEYWORDS	STS; PCR primer; STS sequence; microsatellite DNA; microsatellite marker; sequence tagged site; tetranucleotide repeat.		
SOURCE	Homo sapiens DNA.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 91)		
	Gerken,S.C., Matsunami,N., Plaecke,R., Albertsen,H., Ballard,L., Mellis,R., Lawrence,E., Moore,M., Holik,P.R., Carlson,M., Zhao,X., Robertson,M., Bradley,P., Elser,T., Tingey,A., Lalouel,J.-M. and White,R.		
TITLE	Genetic and physical mapping of simple sequence repeat containing sequence tagged sites from the human genome		
JOURNAL	Unpublished (1994)		
COMMENT	Submitted by: Utah Center for Human Genome Research University of		

[illegible]

Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 103)
AUTHORS Aldridge,F.L.
TITLE Direct Submission
JOURNAL Submitted (12-FEB-1991) F.L. Aldridge, ICI Pharmaceuticals,
Alderley Park, Macclesfield, Cheshire, SK10 4TG, UK
REFERENCE 2 (bases 1 to 103)
AUTHORS Butler,R., Riley,J.H., Ogilvie,D.J., Anand,R., Buxton,J.,
Davies,J., Johnson,K. and Markham,A.F.
TITLE Two sequence-tagged sites defining the ends of a 380 kb YAC clone
from 19q13
JOURNAL Nucleic Acids Res. 19 (17), 4787 (1991)
MEDLINE 91367697
COMMENT See also X57788 for STS 81C8.
FEATURES
source
1..103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="19q13"
/geneLine
/clone.lib="YAC library: ICI"
BASE COUNT 29 a 28 c 23 g 22 t 1 others
ORIGIN

Query Match 0.4%; Score 75; DB 13; Length 103;
Best Local Similarity 84.0%; Pred. No. 0.019; Length 103;
Matches 84; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OY 8133 GCCTGTAATCCAGCATTGGCAGCTGAGCGGAGCATCTTGGAGTCAGAGATT 8192
DB 101 GCGTATTAATTTCGCTGGAGGTCGAGCGGAGTCGATCACTTAAGTCAGAGATT 42

OY 8193 GAGACGAGCTGCGCAACATGTTGAACCCCTGCTCCACT 8232
DB 41 TTGACGAGCTGGCCCAACATGTTGAACCCCTGCTCCACT 2

RESULT 14
LOCUS HUMALCE43 110 bp ss-RNA PRI 15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE43.
ACCESSION M87900.1 GI:174876
VERSION M87900.1 GI:174876
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 110)
AUTHORS Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of
post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
source
1..110
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="NTera2D1"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"
BASE COUNT 27 a 31 c 34 g 18 t
ORIGIN

Query Match 0.4%; Score 74.4; DB 9; Length 110;
Best Local Similarity 80.6%; Pred. No. 0.023;
Matches 87; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

OY 12893 GTTTCGCGCTTTAGCAGAGATGTCATCTCCGACCTGCTGATCCAGCGGCTCGG 12952
DB 110 GTTTCGCTCATTTAGCAGAGATGTCATCTCCGACCTGCTGATCCAGCGGCTCGG 51

OY 12953 CCTCCCAAGTGTGGATTACAGGATGCGCACACAGCCTGCGCG 13000
DB 50 CCTCCCAAGTGTGGATTACAGGATGCGCACACAGCCTGCGCGCGCGCG 3

RESULT 15
LOCUS HUMD1D03M5 108 bp mRNA PRI 04-FEB-1999
DEFINITION Human HepG2 partial cDNA, clone hmd1d03m5.
ACCESSION D16965
VERSION D16965.1 GI:598552
KEYWORDS gene signature.
SOURCE Homo sapiens Male cell_line:HepG2 cDNA to mRNA, clone.lib:Kiservu.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Matoba,R.
TITLE Direct Submission
JOURNAL Submitted (21-JUL-1993) to the DDBJ/EMBL/GenBank databases. Ryo
Matoba, Osaka University, Institute for Molecular and Cellular Bio;
1-3, Yamada-oka, Suita, Osaka 565, Japan
(E-mail:matoba@inherit.imcb.osaka-u.ac.jp,
Tel:81-6-877-5111(ex.3314), Fax:81-6-877-1922)
2 (bases 1 to 108)
REFERENCE 2 (bases 1 to 108)
AUTHORS Matoba,R., Okubo,K., Horii,N., Fukushima,A. and Matsubara,K.
TITLE The addition of 5' coding information to a 3'-directed cDNA library
Improves analysis of gene expression
Gene 146 (2), 199-207 (1994)
COMMENT Submitted (21-Jul-1993) to DDBJ by:
Ryo Matoba
Molecular Microbiology and Genetics Lab.
Research Institute of Innovative Technology for the Earth 9-2
Kizugawada Kizu-cyo,
Soraku-gun, Kyoto
Japan, 619-02
Phone: 07747-5-2308
Fax: 07747-5-2321.
FEATURES
source
1..108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="HepG2"
/clone.lib="Kiservu"
/sex="Male"
BASE COUNT 28 a 23 c 38 g 17 t 2 others
ORIGIN

Query Match 0.4%; Score 74.2; DB 9; Length 108;
Best Local Similarity 85.3%; Pred. No. 0.024; Length 108;
Matches 93; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

OY 13467 GATCTTGCTCATTTGCAACCTGCTGCTGGTTCAGTGAATTCCTGATCAGCCTC 13526
DB 108 GATCTTGCTCATTTGCAACCTGCTGCTGGTTCAGTGAATTCCTGATCAGCCTC 49

OY 13527 CCGAGTAGCTGGGATTACAGGATGATCAGCATGCGCTGGTAAATTTT 13575
DB 48 CTGAGTAGCTGGGATTACA-GCATGCGGCACACACACNCTGCTTTTAT 1

Search completed: June 17, 2000, 10:40:36
Job time: 252120 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2000, 11:43:03 ; Search time 939.94 Seconds

(without alignments)
4791.484 Million cell updates/sec

Title: US-08-852-495c-1_COPY_196000_214000

Perfect score: 18001

Sequence: 1 GATAGGCTGACTTCTACCA.....CCATCCAGCTTCTCCTCCT 18001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 311585 seqs, 125096042 residues

Total number of hits satisfying chosen parameters: 433070

Minimum DB seq length: 10

Maximum DB seq length: 110

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	71.8	0.4	108	1 X12095	Human biallelic po
2	71.4	0.4	108	1 X01323	Allelic ladder, H
3	68.6	0.4	108	1 X12095	Human biallelic po
4	65.6	0.4	100	1 T24892	Human gene signatu
5	65.2	0.4	103	1 T26213	Human gene signatu
6	64	0.4	108	1 T25009	Human gene signatu
7	63.8	0.4	108	1 T26828	Human gene signatu
8	60.2	0.3	100	1 T24892	Human gene signatu
9	58.6	0.3	103	1 T20927	Human gene signatu
10	58.4	0.3	87	1 T21566	Human gene signatu
11	57.4	0.3	85	1 O95218	Simple tandem repe
12	57	0.3	100	1 X12085	Human biallelic po
13	57	0.3	100	1 X12086	Human biallelic po
14	56.8	0.3	93	1 T22572	Human gene signatu
15	56.4	0.3	82	1 T22836	Human gene signatu
16	56.4	0.3	91	1 T25854	Human gene signatu
17	55.6	0.3	100	1 X12087	Human biallelic po
18	55.4	0.3	106	1 V11611	Homo sapiens adult
19	54.8	0.3	103	1 T20927	Human gene signatu
20	54.2	0.3	93	1 T24259	Human gene signatu
21	54	0.3	109	1 T23895	Human gene signatu
22	53.8	0.3	100	1 X12087	Human biallelic po
23	53.8	0.3	100	1 X12085	Human biallelic po
24	53.8	0.3	100	1 X12086	Human biallelic po
25	52.6	0.3	95	1 T23131	Human gene signatu
26	52.6	0.3	97	1 T26728	Human gene signatu
27	52.4	0.3	87	1 T21566	Human gene signatu
28	52.2	0.3	107	1 T20373	Human gene signatu
29	52.2	0.3	109	1 T23895	Human gene signatu
30	52	0.3	69	1 O29016	Probe to internal
31	52	0.3	69	1 O29016	Probe to internal
32	52	0.3	110	1 T26288	Human gene signatu
33	51.6	0.3	99	1 T23728	Human gene signatu
34	51	0.3	70	1 N60231	Normal chromosome

35	51	0.3	91	1 T25854	Human gene signatu
36	51	0.3	99	1 T20931	Human gene signatu
37	50.8	0.3	85	1 T25730	Human gene signatu
38	50.6	0.3	110	1 T26288	Human gene signatu
39	50.4	0.3	92	1 V11595	Homo sapiens adult
40	50.4	0.3	92	1 V61480	Homo sapiens adult
41	50	0.3	108	1 T26828	Human gene signatu
42	49.6	0.3	93	1 T25688	Human gene signatu
43	49.4	0.3	97	1 T26728	Human gene signatu
44	49	0.3	65	1 T25588	Human gene signatu
45	48.4	0.3	75	1 T22841	Human gene signatu

ALIGNMENTS

RESULT 1
ID X12095
AC X12095;
DE 30-MAR-1999 (first entry)
KW Polymorphism; biallelic; human; forensic; paternity testing; disease;
KW detection; phenotypic typing; characteristic; infection; hereditary;
KW autoimmune disease; cancer; inflammation; drug; therapy; medication;
KW treatment; marker; ss.
OS Homo sapiens.
PN W09820165-A2.
PD 14-MAY-1998.
PF 05-NOV-1997; U20313.
PR 06-NOV-1996; US-030455.
PA (WHED) WHITEHEAD INST BIOMEDICAL RES.
PI Hudson T, Lander ES, Wang D;
DR WPI; 98-286974/25.
PT New isolated nucleic acid segments from the human genome - used for
PT determining polymorphic forms for use in e.g. forensics, paternity
PT testing or phenotypic typing for disease
PS Claim 1: Page 219; 310pp; English.
CC X10269-X12937 are human DNA fragments which contain biallelic polymorphic
CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-IUB ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Lesch-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberous sclerosis, hereditary
CC hemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.
SQ Sequence 108 BP; 19 A; 23 C; 28 G; 37 T;

Query Match 0.4%; Score 71.8; DB 1; Length 108;
Best Local Similarity 84.9%; Pred. No. 0.025; Mismatches 13; Indels 0; Gaps 0;
Matches 79; Conservative 1; Mismatches 13; Indels 0; Gaps 0;
QY 13575 TGTATTTTGTAGAGATGGGGTTTCACATGTTGACAGGCTGCTCAAACTCTGAC 13634
DB 1 TGTCTTTTGTAGAGATGGAGTTTCCTGTGTGGCCAGAGATGCTCAACTCTGAC 60
QY 13635 CTCAGTGTATCCACTCTGCTTACGCTCCCAAA 13667
DB 61 TTCAAGTATCCGTCTGCTTGGCTCCCAAA 93

[illegible]

CC markers which have been isolated using the primers represented in
CC X09121-X10268. The base occupying the polymorphic site is indicated by
CC the appropriate IUPAC-10B ambiguity code. These fragments can be used in
CC methods for determining polymorphic forms in an individual for use in
CC e.g. Forensics, paternity testing or for phenotypic typing for diseases
CC such as agammaglobulinemia, diabetes insipidus, Leech-Nyhan syndrome,
CC muscular dystrophy, Wiskott-Aldrich syndrome, Fabry's disease, familial
CC hypercholesterolemia, polycystic kidney disease, hereditary
CC spherocytosis, von Willebrand's disease, tuberos sclerosis, hereditary
CC hemorrhagic telangiectasia, familial colonic polyposis, Ehlers-Danlos
CC syndrome, osteogenesis imperfecta, acute intermittent porphyria,
CC autoimmune diseases, inflammation, cancer, diseases of the nervous
CC system, infection by pathogenic microorganisms, and characteristics such
CC as longevity, appearance (e.g. baldness, obesity), strength, speed,
CC endurance, fertility, and susceptibility or receptivity to particular
CC drugs or therapeutic treatments. The isolated polymorphic nucleic acid
CC segments can also be used to produce medicaments for the treatment or
CC prophylaxis of such diseases.

CC Sequence 108 BP, 19 A, 23 C, 28 G, 37 T;

Query Match	Similarity	0.4%	Score 68.6	DB 1	Length 108
Best Local	88%	Conservative	1	Mismatches 15	Indels 1
Matches	88	Conservative	1	Mismatches 15	Indels 1
Gaps	1				
Query	4189	TAATCCACAGAC	-TTTGGAGGCCCGACGAGCATCATCTAGGTCGAAATTCACAGA	4247	
Db	105	TAATCCACAGCATTTTGGAGGCCCAACGACGATCTACTGAGTTCAGGATTCGACA	46		
Qy	4248	CCAGCTGGGCCAAAATGCTGAACCTCATCTCTACTAAATAATCA	4292		
Db	45	CCATCTGGCCCAACAYAGGAAACCTCATCTCTACAAAAAACA	1		
RESULT	4				
T24892/c					
ID	T24892	standard; cDNA to mRNA; 100 BP.			
AC	T24892				
DT	05-NOV-1996	(first entry)			
DE	Human gene signature HUMGS06998.				
KW	Gene signature; messenger RNA; mRNA; relative abundance; frequency;				
RV	human; cloning; mapping; non-biased library; diagnosis; detection;				
OS	cell typing; abnormal cell function; ss.				
OS	Homo sapiens.				
PN	W09514772-A1.				
PD	01-JUN-1995.				
PF	11-NOV-1994; J01916.				
PR	12-NOV-1993; JP-355504.				
PA	(MATS/) MATSUBARA K.				
PA	(OKUB/) OKUBO K.				
P1	Matsubara K, Okubo K;				
DR	WPI; 95-206931/27.				
PT	Identifying gene signatures in 3'-directed human cDNA library - e.g.				
PT	for diagnosis of abnormal cell function, by preparing cDNA that				
PT	reflects relative abundance of corresp. mRNA in specific human				
PT	tissues				
P5	Claim 1: Page 1720; 2245bp; Japanese.				
CC	A single-stranded DNA (or its complementary strand or the corresp.				
CC	double-stranded DNA) which comprises one of the 7837 "GS" sequences				
CC	given in T19001-T26837 and which is able to hybridise to part of				
CC	human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)				
CC	sequences were obtained from 3'-directed cDNA libraries prepared				
CC	from various human tissues; synthesis of cDNA was initiated from the				
CC	3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-				
CC	untranslated sequence is unique to a particular mRNA species, almost				
CC	all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library				
CC	is constructed so as to reflect accurately the relative abundance of				
CC	different mRNAs in the particular tissue from which it was derived.				
CC	The appearance frequency of a given GS in a cDNA library can be				
CC	determined (esp. using primers and probes derived from the GS				
CC	sequences) as a means of diagnosing abnormal cell function or for				
CC	recognising different cell types.				
Sequence	100 BP;	28 A;	22 C;	25 G;	22 T;


```
Query Match      0.4%; Score 65.6; DB 1; Length 100;
Best Local Similarity 77.8%; Pred. No. 0.18; Indels 0; Gaps 0;
Matches 77; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 12707 TTTTGTGTTTCTGACAGCTGTTGCTCTGCTCAGAGGCTGAGTGGTGCATC 12766
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 100 TTTTGTGTTTCTGACAGAGTGTCTACTCTCTACCCAGCGGAGTGCAGTGCATC 41

QY 12767 TCGGCTCACTGCAACCTGCTGCCGCCGGGTTCAGTGAT 12805
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 40 TCAGCTATGTCAAATTCCTCTCCAGGTTCAAGCGAT 2

RESULT 5
T26213/c
ID T26213 standard; cDNA to mRNA; 103 BP.
AC T26213;
DT 13-NOV-1996 (first entry)
DE Human gene signature HUMGS08452.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN MO9514772-A1.
PD 01-JUN-1995.
PE 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K. Okubo K.
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1: Page 2029; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 103 BP; 33 A; 21 C; 25 G; 23 T;
```

```
DT 07-NOV-1996 (first entry)
DE Human gene signature HUMGS07131.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN MO9514772-A1.
PD 01-JUN-1995.
PE 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K. Okubo K.
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1: Page 1748; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 108 BP; 34 A; 31 C; 26 G; 15 T;
```

```
Query Match      0.4%; Score 64; DB 1; Length 108;
Best Local Similarity 74.5%; Pred. No. 0.3; Indels 0; Gaps 0;
Matches 79; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 15026 TTTGTTGTTGTTGTTTGTGATAGACTCTCTCTCATTTCAATTCAGGCTGAGTGCATGG 15085
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 107 TTGATGTTGTTGTTGTTTCAACAGGCTCTTCTCTCTCACTGAGTGCATTCAGTGCATGG 48

QY 15086 CATGATTCAGCTCAGCTGACAGCTCCGCTCCGCGTTCAAGAGAT 15131
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 47 CGTGACCATGCTCAGCTGACAGCTTGGCCTCATGAGTGCATGCAGGCGAT 2

RESULT 7
T26828
ID T26828 standard; cDNA to mRNA; 108 BP.
AC T26828;
DT 14-NOV-1996 (first entry)
DE Human gene signature HUMGS09078.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN MO9514772-A1.
PD 01-JUN-1995.
PE 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K. Okubo K.
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1: Page 2182; 2245pp; Japanese.
```

CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 108 BP; 18 A; 33 C; 23 G; 28 T;

Query Match 0.4%; Score 63.8; DB 1; Length 108;
Best Local Similarity 88.3%; Pred. No. 0.32;
Matches 68; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 12921 GATCTCTGACCTCGATGACACCGGCTGCTCCCAAGCTGGATTACAGGCAT 12980
|||||
DB 1 GATCTCTGACCTCGATGACCGCCGCTGCTCCCAAGCTGGATTACAGGCAT 60
OY 12981 GGGCAGCAGCGCTGGC 12997
|
DB 61 GAGCCACACGCGCCGCG 77

RESULT 8
T24892
ID T24892 standard; cDNA to mRNA; 100 BP.
AC T24892; (first entry)
DE Human gene signature HUMGS06598.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PE 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues.
PS Claim 1. Page 1720; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 100 BP; 28 A; 22 C; 25 G; 22 T;

Query Match 0.3%; Score 60.2; DB 1; Length 100;

Best Local Similarity 74.0%; Pred. No. 1;
Matches 74; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

OY 4357 GATCACTTGAACTGGAGGAGAGATGTCATGACTGAGATCAATGCCATGATCCA 4416
|||||
DB 1 GATCACTTGAACTGGAGGAGAGATGTCATGACTGAGATCAATGCCATGATCCA 60
OY 4417 GCCTGGGAG 4456
|||||
DB 61 GCCTGGGAG 100

RESULT 9
T20927/c
ID T20927 standard; cDNA to mRNA; 103 BP.
AC T20927; (first entry)
DE Human gene signature HUMGS02180.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-A1.
PD 01-JUN-1995.
PE 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues.
PS Claim 1. Page 758-759; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 103 BP; 22 A; 27 C; 21 G; 31 T;

Query Match 0.3%; Score 58.6; DB 1; Length 103;
Best Local Similarity 73.7%; Pred. No. 1.7;
Matches 73; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

OY 4262 ATGTGTAACCTGATCTCTAAATAACAAAGAGAGAGAGAGAGAGAGAGAG 4321
|||||
DB 99 ATGTGTAACCTGATCTCTCTAAATAACAAAGAGAGAGAGAGAGAGAGAGAG 40
OY 4322 TGTATATCCAGTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 4360
|||||
DB 39 TGTATATCCAGTACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1

RESULT 10
T21566/c
ID T21566 standard; cDNA to mRNA; 87 BP.
AC T21566; (first entry)
DE Human gene signature HUMGS02944.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;

CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 82 BP; 26 A; 25 C; 14 G; 17 T;

Query Match 0.3%; Score 56.4; DB 1; Length 82;

Best Local Similarity 80.5%; Pred. No. 3.4; Mismatches 0; Gaps 0;

Matches 66; Conservative 0; Indels 16; Gaps 0;

OY 4542 TTTAAGTTCCCTGGGTACAGTACAGATGTCAGGTTGTTACATAGTAACGTGTGCC 4601

Db 82 TTTAAGTTCTGGGGTACATGCAATGATGTGCAAGTTGTTATGACAGTAATGTGTGCC 23

OY 4602 ATGGTGATTTGCTGCACCTATC 4623

Db 22 ACGGAGGTTGCTCTCCACATC 1

Search completed: June 17, 2000, 11:43:23
Job time: 254714 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2000, 02:34:25 ; Search time 13753.1 Seconds
(without alignments)
5305.117 Million cell updates/sec

Title: US-08-852-495C-1_COPY_196000_214000

Perfect score: 18001

Sequence: 1 GATAGCCTCACTTCTACCA.....CCATCAGCTTCTCACTCTCT 18001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 202611650 residues

Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10

Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: gb_est1:*
21: gb_est2:*
22: gb_est3:*
23: gb_est4:*
24: gb_est5:*
25: gb_est6:*
26: gb_est7:*
27: gb_est8:*
28: gb_est9:*
29: gb_est10:*
30: gb_est11:*
31: gb_est12:*
32: gb_est13:*
33: gb_est14:*
34: gb_est15:*
35: gb_est16:*
36: gb_est17:*
37: gb_est18:*
38: gb_est19:*
39: gb_est20:*
40: gb_est21:*
41: gb_est22:*
42: gb_est23:*
43: gb_est24:*
44: gb_est25:*

45: gb_est26:*
46: gb_est27:*
47: gb_est28:*
48: gb_est29:*
49: gb_est30:*
50: gb_est31:*
51: gb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: gb_est33:*
60: gb_est34:*
61: gb_est35:*
62: gb_est36:*
63: gb_est37:*
64: gb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: gb_est39:*
70: gb_est40:*
71: gb_est41:*
72: gb_est42:*
73: gb_est43:*
74: gb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: gb_est45:*
80: gb_est46:*
81: gb_est47:*
82: gb_gsa1:*
83: gb_gsa2:*
84: gb_gsa3:*
85: gb_gsa4:*
86: em_gsa1:*
87: em_gsa2:*
88: em_gsa3:*
89: em_gsa4:*
90: gb_gsa5:*
91: gb_gsa6:*
92: gb_gsa7:*
93: gb_gsa8:*
94: gb_gsa9:*
95: em_gsa5:*
96: em_gsa6:*
97: em_gsa7:*
98: em_gsa8:*
99: em_gsa9:*
100: em_gsa10:*
101: em_gsa11:*
102: gb_gsa10:*
103: gb_gsa11:*
104: em_gsa12:*
105: gb_gsa12:*
106: gb_gsa13:*
107: gb_gsa14:*
108: gb_gsa15:*
109: gb_gsa16:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result

Query

SUMMARIES

No.	Score	Match	Length	DB	ID	Description
1	93	0.5	109	30	AA243009	AA243009 zr25h02.s
2	91.6	0.5	106	37	AA703692	AA703692 ag81a10.r
3	91	0.5	107	35	AA565533	AA565533 nk42d11.s
4	90.2	0.5	103	84	B48912	B48912 RPII11-4A12
5	90	0.5	110	108	AQ386882	AQ386882 RPII11-13
6	88.6	0.5	103	106	AQ535244	AQ535244 RPII11-13
7	87.6	0.5	110	106	AQ386882	AQ386882 RPII11-13
8	87.4	0.5	106	108	AQ544957	AQ544957 CTRBI-E1
9	86.8	0.5	110	39	AA897366	AA897366 am06h02.s
10	86.6	0.5	103	38	AA807640	AA807640 nx08b03.s
11	86.6	0.5	109	84	B17434	B17434 34SK2.TVB C
12	86.6	0.5	109	84	B17434	B17434 34SK2.TVB C
13	86.6	0.5	110	30	AA244245	AA244245 nc07a04.s
14	86.6	0.5	110	94	AQ0003188	AQ0003188 RPII11-1D
15	86.2	0.5	107	35	AA565533	AA565533 nk42b11.s
16	86.2	0.5	108	84	B65160	B65160 CTR-HSP-201
17	86	0.5	105	109	AQ637292	AQ637292 RPII11-4
18	85.4	0.5	103	94	AQ028649	AQ028649 CTR-HSP-2
19	85.4	0.5	103	108	AQ535244	AQ535244 RPII11-3
20	85	0.5	106	30	AA250812	AA250812 zs06a05.s
21	84.4	0.5	102	36	AA654562	AA654562 nt75f10.s
22	83.6	0.5	106	105	AQ264176	AQ264176 CTRBI-E1
23	83.4	0.5	109	105	AQ265749	AQ265749 CTRBI-E1
24	83	0.5	107	35	AA583252	AA583252 nm41e04.s
25	82	0.5	109	94	AQ029690	AQ029690 RPII11-41
26	82	0.5	103	108	AQ582186	AQ582186 RPII11-4
27	82	0.5	106	63	AT991750	AT991750 wt48e01.x
28	82	0.5	106	63	AT991750	AT991750 wt48e01.x
29	81.6	0.5	107	24	H67040	H67040 yu68c01.r1
30	81.6	0.5	108	35	AA594869	AA594869 no21e02.s
31	81.2	0.5	102	84	B48088	B48088 RPII11-4N6
32	81.2	0.5	104	108	AQ544583	AQ544583 CTRBI-E1
33	81	0.4	105	28	AA078003	AA078003 7H12D08 C
34	80.8	0.4	107	106	AQ412658	AQ412658 RPII11-2
35	80.8	0.4	108	84	B32951	B32951 HS-1016-A1-
36	80.8	0.4	109	103	AQ0200347	AQ0200347 RPII11-43
37	80.6	0.4	103	35	AA570438	AA570438 nk63902.s
38	80.6	0.4	103	108	AQ534922	AQ534922 RPII11-13
39	80.6	0.4	108	84	B15423	B15423 345B10.TV C
40	80.4	0.4	106	38	AA812141	AA812141 ob4h02.s
41	80.4	0.4	109	24	N25299	N25299 yw32c09.s1
42	80.2	0.4	106	30	AA250812	AA250812 zs06a05.s
43	80.2	0.4	109	94	AQ028426	AQ028426 CTR-HSP-2
44	80.2	0.4	109	105	AQ265749	AQ265749 CTRBI-E1
45	80	0.4	97	39	AA837701	AA837701 oe06c02.s

ALIGNMENTS

RESULT 1
AA243009 109 bp mRNA EST 11-MAR-1998
LOCUS zr25h02.s1 Stratagene NM2 neuronal precursor 937230 Homo sapiens
DEFINITION cdna clone IMAGE:664467 3' similar to contains Alu repetitive
element;contains element LTR1 repetitive element ; mRNA sequence.
ACCESSION AA243009
VERSION AA243009.1 GI:1873869
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 109)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kritman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
Unpublished (1997)
On Dec 3, 1996 this sequence version replaced gi:1126869.

FEATURES
source
1..109
/organism="Homo sapiens"
/db_xref="GDB:5426481"
/db_xref="taxon:9606"
/clone="IMAGE:664467"
/clone_lib="Stratagene NM2 neuronal precursor 937230"
/issue_type="neuroepithelial cells"
/dev_stage="Ntera-2 neuroepithelial cells"
/note="Organ: brain; Vector: pBluescript SK-; Site:1:
EcoRI; Site:2: XhoI; Cloned unidirectionally. Primer:
Oligo dr. Uninduced, exponentially growing neuroepithelial
cells (Ntera-2/cl.D1). Average insert size: 1.0 kb;
Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG
3' -3' adaptor sequence: 5' CTCAGGTTTCTTTTCTTTT 3'."

BASE COUNT 19 a 30 c 30 g 30 t
ORIGIN

Query Match 0.5%; Score 93; DB 30; Length 109;
Best Local Similarity 90.8%; Pred. 0.062;
Matches 99; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 12873 GATTTTATAGAGATGAGTTTCCGCGTTTACGACAGATGTCGATCTCCGACC 12932
|||||
DB 1 GATTTTATAGAGATGAGCGGGGTTTACCGGTTTACCGATGTCGATCTCCGACC 60
OY 12933 TCGTGATCCGCGCCGCGCCGCGCCGCGCGGATGCTGGATTCAGCGCATG 12981
|||||
DB 61 TCGTGATCCGCGCCGCGCCGCGCGCGGATGCTGGATTCAGCGCATG 109
|||||

RESULT 2
AA703692 106 bp mRNA EST 24-DEC-1997
LOCUS ag81a10.r1 Stratagene hmt neuron (#937233) Homo sapiens cdna clone
DEFINITION IMAGE:1140858 5' similar to contains Alu repetitive element;; mRNA
sequence.
ACCESSION AA703692
VERSION AA703692.1 GI:2713610
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 106)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Kritman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
Unpublished (1997)
On Sep 12, 1996 this sequence version replaced gi:1397630.
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.

Seq primer: -28m13 rev1 ET from Amersham
High quality sequence stop: 53.
Location/Qualifiers

FEATURES

source

1..106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1140858"
/clone_lib="Stratagene hMT neuron (#937233)"
/dev_stage="hMT neurons"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt.
Differentiated, post mitotic hMT neurons. Average insert size: 1.5 kb. Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor sequence: 5' CTCGAGTGTGTGTGTGTGT 3'"

BASE COUNT

19 a 29 c 29 g 29 t

Query Match 0.5%; Score 91.6; DB 37; Length 106;
Best Local Similarity 91.5%; Pred. No. 0.091;
Matches 97; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 12876 TTTTAGTAGAGATGGATGGCGGTGTAGCGAGGATGCTCGATCTCTGACCTCG 12935
|||||
Db 1 TTTTGTAGAGAGCGAGGTTCACCGGTTCAGAGATGCTCTCGATCTCTGACCTCG 60
QY 12936 TGAATCCAGCGGCTCGGCTCCCAAGATGCTGATTCAGCATG 12981
|||||
Db 61 TGAATCGCCCGCTCAGCCTCCCAAGATGCTGATTCAGCATG 106

RESULT 3

AA565533 107 bp mRNA EST 08-SEP-1997
LOCUS nK42b11.s1 NCI-CGAP GC2 Homo sapiens cDNA IMAGE:1016157 3'
DEFINITION similar to contains Alu repetitive element; mRNA sequence.
ACCESSION AA565533
VERSION AA565533.1 GI:2337172
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 107)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Sep 12, 1996 this sequence version replaced gi:1393355.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Stratagene, Inc., David B. Kitzman,
Ph.D.

CDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www.bio.lnl.gov/db/rdp/image/image.html

JOURNAL COMMENT

FEATURES

source

Insert Length: 1661 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 87.
Location/Qualifiers

1..107
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1016157"
/clone_lib="NCI-CGAP_GC2"

/tissue_type="germ cell tumor"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Bulk germ cell tumor. 5' adaptor sequence: 5' GAATTCGGCAGCAG 3' 3' adaptor sequence: 5' CTCGAGTGTGTGTGTGTGT 3'
Average insert size: 1.2 kb."

BASE COUNT

22 a 34 c 26 g 25 t

Query Match 0.5%; Score 91; DB 35; Length 107;
Best Local Similarity 90.7%; Pred. No. 0.11;
Matches 97; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 12746 CTGAGTGCAGTGGTGGCTCTCGGCTCACTGCAGACCTCTGCTCCGGGTCAAGTAT 12805
|||||
Db 1 CTGGAGTGCAGTGGTGGCTCAATCTCAGTCACTGCAGACCTCTGCTCCAGGTTCAAGTAT 60
QY 12806 TCTCTGCTCTCAGCTCTCCGAGTACGTTGACTACAGCAGCACACCA 12852
|||||
Db 61 TCTCTGCTCTCAGCTCTCCGAGTACGTTGACTACAGCAGCACACCA 107

RESULT 4

B48914/c 103 bp DNA GSS 08-APR-1999
LOCUS RPCI11-4A12.JP RPCI-11 Homo sapiens genomic clone RPCI-11-4A12,
DEFINITION genomic survey sequence.
ACCESSION B48914
VERSION B48914.1 GI:2601151
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 103)
Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wilde,C., de Jong,P. and
Venter,J.C.
Use of BAC End Sequences for Sequence-Ready Map Building
Unpublished (1997)
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamad@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from
Research Genetics (inforesgen.com). BAC end search page:
http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html
Seq primer: Sp6
Class: BAC ends.

JOURNAL COMMENT

Unpublished (1997)
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamad@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from
Research Genetics (inforesgen.com). BAC end search page:
http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html
Seq primer: Sp6
Class: BAC ends.

FEATURES

source

Location/Qualifiers

1..103
/organism="Homo sapiens"
/db_xref="GDB:7501163"
/db_xref="taxon:9606"
/clone="RPCI-11-4A12"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"

BASE COUNT

30 a 28 c 30 g 15 t

Query Match 0.5%; Score 90.2; DB 84; Length 103;

[illegible]

RESULT	5
AQ386882/c	
LOCUS	AQ386882 110 bp DNA GSS 21-MAY-1999
DEFINITION	RCCII11-14I4.TV RPCI-11 Homo sapiens genomic clone RPCI-11-13A14.
ACCESSION	AQ386882 genomic survey sequence.
VERSION	AQ386882.1 GI:4357905
KEYWORDS	GSS.
SOURCE	human.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 110)	Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.	Use of BAC End Sequences from Library RPCI-11 for Sequence-Map Building	Unpublished (1997)	
			Other-SSS: RPC11-13414.TX	

JOURNAL
COMMENT

Unpublished (1997)
Other GSSs: Rpci11-13414..TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel.: 301 838 0200
Fax: 301 838 0208
Email: hbeetlgr.org

Clones are derived from the human BAC library RCI1-11. For BAC library availability, please contact Pieter de Jong (pieter@ejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genetics (inforesgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seq primer: T7
Class: BAC ends.

```

FEATURES
  SOURCE
    1..110
      Location/Qualifiers
        /organism="Homo sapiens"
        /db_xref="GDB:7551267"
        /db_xref="taxon:9606"
        /clone="RPC1-11-134I4"
        /clone_1lb="RPC1-11"
        /sex="Male"
        /cell_type="Lymphocytes"
        /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
        RPc111 Human Male BAC Library"
  BASE COUNT
    26 a      26 c      38 g      20 t
  ORIGIN

```

Query Match	0.5%	Score 90;	DB 106;	Length 110;
Best Local Similarity	90.6%;	Pred. NO. 0.14;		
Matches 96;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0

Oy 1354 GGGTTTCACCATGTTGACCAGCGGTGGTCTCAACTCCTGACCTCAAGTGATCCACCTGCC 13653
|||||
|||||
Db 110 GGGTTTCACCATGTTGTCACAGCGGTGGTCTGAACCTCTGACCTCAAGCATCACCTGCC 51
|||||
|||||

QY 13654 TTAGCCTCCCAAATGCTGGACTACAGCGGTGAGCCACTGCACCC 13699
| | | | | | | | | | | | | | | | | | | | | |
Db 50 TCAGCCTCCCAAGTACTTGATTCACAGCGGTGAGCCACTGCTCCC 5

RESULT	6	
A0535244		
LOCUS	103 bp	DNA
DEFINITION	RPC1-11-31/H22, IV RPC1-11 Homo sapiens genomic clone	GSS 18-MAY-1999
ACCESSION	RPC1-11-31/H22, genomic survey sequence.	
	A0535244	

VERSION	A0535244.1	GI:4840394
KEYWORDS	GSS.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 103)	Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and Venter, J.C.	Use of PAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building	Unpublished (1997)	
	Contact: Shaying, Zhao, William Nierman, Mark Adams			

Email: hbee@igf.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPC Resources (<http://bacpc.med.buffalo.edu/ordering>) or from Research Genet cs (info.genetcs.com). BAC end search page: http://www.cigr.org/cdb/humgen/bac_end_search.html
Seq primer: 17
Class: BAC ends.

```

FEATURES
SOURCE
    location/Qualifiers
    1. .103
        /organism="Homo sapiens"
        /db_xref="GDB:762153"
        /db_xref="taxon:9606"
        /clone="RPC1-11-317H22"
        /clone_11b="RPC1-11"
        /sex="Male"
        /cell_type="Lymphocytes"
        /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
        Rpecl1 Human Male BAC Library"
BASE COUNT
    31 a      27 c      27 g      18 t
ORIGIN

```

Query Match	0.5%	Score 88.6	DB 108	Length 103
Best Local Similarity	91.3%	Pred. No. 0.2		
Matches 94	Conservative 0	Mismatches 9	Indels 0	Gaps 0

QY	8203	TGGGCAACATGCTGAACCCCTGCTCCACATAAAAATACAAAA	8245
Db	1	CCACACATTTGGGGAGCCCAAGACGGGGAGATCACTTGGAGTCCAGGATTCGAGACAGCC	60
QY	8143	CCACACATTTGGGCAAGCGCTGAGGGGGGAGATCACTTGGAGTCCAGGATTCGAGACAGCC	8203
Db	61	TGGGCAACATGCTGAACCCCGTCTTCGCTATATAATACAAAA	103

LOCUS	110 bp	DNA	GSS	21-MAY-1999
DEFINITION	RPC111-13414, TV RPC1-11	Homo sapiens genomic clone RPC1-11-13414,		
DESCRIPTION	genomic survey sequence.			
ACCESSION	U036693			

KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia.

REFERENCE	Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 110) Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter,J.C.
TITLE	User of BAC End Sequences from Llibrary RPCI-11 for Sequence-Ready Map Building Unpublished (1997)
JOURNAL	Other.GSSS: RPCI11-13414.TJ
COMMENT	Contact: Shaying Zhao, William Nierman, Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850 Tel: 301 838 0200 Fax: 301 838 0208 Email: hbe@tigr.org Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@long.med.buffalo.edu). Clones may be purchased from BACPac Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (http://inforesgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html Seq primer: 77 Class: BAC ends.
FEATURES	Location/Qualifiers 1..110 /organism="Homo sapiens" /db_xref="GDB:7551267" /db_xref="taxon:9606" /clone="RPCI-11-13414" /clone_1lb="RPCI-11" /sex="Male" /cell_type="Lymphocytes" /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI; RPCI11 Human Male BAC Library"
BASE COUNT	26 a 26 c 38 g 20 t
ORIGIN	
Query Match	0.5%; Score 87.6; DB 106; Length 110;
Best Local Similarity	87.3%; Pred. No. 0.26;
Matches	96; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
OY	8113 GCCCGGCCCTGGTGCCTCACGCCGTGAATCCAGACACTTTGGCAGGCTGAGCGGGCAGA 8172 Db 1 GCCCGGAGCAATGCGCTCACGCCGTGAATCCAACTTGCGAGAGCTGAJGACAGGTGA 60
OY	8173 TCACCTGAGGTCAGAGATTGAGACAGCGCTGCCAACAATGTGTGAACC 8222 Db 61 TCGCTTAGGTCAGAGATTGAGACAGCGCTGCCAACAATGTGTGAACC 110
RESULT 8	
LOCUS	AO544957 106 bp DNA GSS 28-MAY-1999
DEFINITION	CITBI-EI-2629N2.1F CITBI-EI Homo sapiens genomic clone 2629N2, genomic survey sequence.
ACCESSION	AO544957
VERSION	AO544957.1 GI:4903683
KEYWORDS	GSS.
SOURCE	human. Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and Venter,J.C.
AUTHORS	User of BAC End Sequences from CalTech Libraries for Sequence-Ready Map Building Unpublished (1997)
TITLE	Contact: Shaying Zhao, William Nierman, Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850
JOURNAL	
COMMENT	

	Tel: 301 838 0200
	Fax: 301 838 0208
	Email: hbe@tigr.org
	Clones are available from Research Genetics (info@resgen.com). BAC
	end search page:
	http://www.tigr.org/tcdb/hungen/bac_end_search/bac_end_search.html.
	Seq primer: M13-21
	Class: BAC ends.
FEATURES	
source	Location/Qualifiers
	1..106
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="2629N2"
	/clone_11b="CITBI-EI"
	/sex="male"
	/cell_type="sperm"
	/note="Vector: pBeloBAC11; Site_1: EcoRI; site_2: EcoRI;
	Caltch Human BAC Library P"
BASE COUNT	26 a 26 c 36 g 18 t
ORIGIN	
Query Match	0.5%; Score 87.4; DB 108; Length 106;
Best Local Similarity	89.5%; Pred No 0.28;
Matches	94; Conservative 0; Mismatches 11; Indels 0; Gaps 0
OY	12894 TTTCCGCGTGTAGCCAGATGGTCATCTCATCTGCATCTGCATGCCGCGCTCGGC 12953
DB	106 TTTCACCGTGTTAGCCAGATGGTCATCTCATCTGCATCTGCATGCCGCGCTCGGC 47
OY	12954 CTCGCCAAGTCTGGGATTACAGGCAAGTGGGCCACCACCGCTGGCC 12998
DB	46 CTCGCCAAGTCTGGGATTACAGGCTTGGGCCCTCCACGCCGCGCC 2
RESULT	9
AA897366	
LOCUS	AA897366 110 bp mRNA EST 04-JAN-1999
DEFINITION	am06h02.s1 Soares_NFL_T_GBC.S1 Homo sapiens CDNA clone
	IMAGE:146067 3' similar to contains Alu repetitive element;; mRNA
	sequence.
ACCESSION	AA897366
VERSION	AA897366.1 GI:3033986
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
	Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE	1 (bases 1 to 110)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL	Tumor Gene Index
COMMENT	Unpublished (1997)
	On Jan 19, 1998 this sequence version replaced gi:2150764.
	Contact: Robert Strausberg, Ph.D.
	Tel: (301) 496-1550
	Email: Robert.Strausberg@nih.gov
	This clone is available royalty-free through LNLN ; contact the
	IMGE Consortium (info@image.lnl.gov) for further information.
	Insert length: 834 Std Error: 0.00
	Seq primer: -40m13 fwd. ET from Amersham
	High quality sequence stop: 63.
FEATURES	
source	Location/Qualifiers
	1..110
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone_11b="Soares_NFL_T_GBC.S1"
	/lab_host="DH10B"
	/note="Organ: pooled: Vector: p7r3d-Pac (Pharmacia) with
	a modified polylinker, site_1: Not I, site_2: Eco RI;
	Equal amounts of plasmid DNA from three normalized
	libraries (fetal lung NBHL19W, testis NH7, and B-cell

NCI-CGAP (CBI) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 297480-302087, 682632-687239, 726408-728711, and 729096-731399. Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 22 a 27 c 29 g 32 t

ORIGIN

Query Match 0.5%; Score 86.8; DB 39; Length 110;
Best Local Similarity 88.7%; Pred. No. 0.32;
Matches 94; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

OY 12876 TTTTATGAGAGATGAGATTTCGCGGTGTAGCCAGATGCTTCGATCTCTGACCTCG 12935
|||||
DB 5 TTTTATGAGAGATGAGATTTCGCGGTGTAGCCAGATGCTTCGATCTCTGACCTCA 64
OY 12936 TGATCCAGCGCTCGGCTCCCAAGTGTGGATTTACAGGCAATG 12981
|||||
DB 65 TGATCCAGCGCTCGGCTCCCAAGTGTGGATTTACAGGCGTG 110

RESULT 10
AA807640 103 bp mRNA EST 05-MAR-1998
DEFINITION nx08b05.s1 NCI-CGAP_GC3 Homo sapiens cDNA clone IMAGE:1255473 3'
ACCESSION AA807640
VERSION AA807640.1 GI:2877108
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2151346.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskalko, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
CNA Library Preparation: M. Bento Soares, Ph.D.
DNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNW at:
www-bio.llnl.gov/dbfp/image/image.html

Insert Length: 774 Std Error: 0.00
Seg primer: -40m3 fwd. ET from Amersham
High quality sequence stop: 87.
Location/Qualifiers
1. 103

FEATURES
SOURCE
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1255473"
/clone_lib="NCI-CGAP_GC3"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT73
vector. Library is not normalized. Library was

BASE COUNT 19 a 27 c 30 g 27 t
ORIGIN constructed by Bento Soares and M. Fatima Bonaldo.

Query Match 0.5%; Score 86.6; DB 38; Length 103;
Best Local Similarity 91.1%; Pred. No. 0.35;
Matches 92; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 12881 AGTAGAGATGAGTTTGGCGGTGTAGCCAGATGATGTCGATCTCGACCTGATG 12940
|||||
DB 2 AGTAGAGATGAGTTTGGCGGTGTAGCCAGATGATGTCGATCTCGACCTGATG 61
OY 12941 CACCGGCTTCGCGCTCCCAAGTGTGGATTTACAGGCAATG 12981
|||||
DB 62 CGTCACCTCGGCTCCCAAGTGTGGATTTACAGGCTGTG 102

RESULT 11
B17434 109 bp DNA GSS 04-JUN-1998
DEFINITION B17434.1 CTR978SK1 Homo sapiens genomic clone A-345K02, genomic
survey sequence.
ACCESSION B17434
VERSION B17434.1 GI:2125183
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 109)
AUTHORS Adams, M.D., Kelley, J.M., Rounsley, S.R. and Venter, J.C.
TITLE Use of a BAC End Sequence Database for Sequence-Ready Map Building
JOURNAL Other GSS: 345K02.TP 345K02.TPB
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Seg primer: T7
Class: BAC ends.

FEATURES
SOURCE
Location/Qualifiers
1. 109

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="A-345K02"
/clone_lib="CTR978SK1"
/sex="Female"
/cell_type="Fibroblast"
/note="Vector: pBAC108L, site_1: HindIII; site_2: HindIII;
Caltech Human BAC Library A1"
BASE COUNT 24 a 30 c 31 g 24 t
ORIGIN

Query Match 0.5%; Score 86.6; DB 84; Length 109;
Best Local Similarity 87.2%; Pred. No. 0.34;
Matches 95; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

OY 8125 TGCTACACCTGTATATCCAGCACTTTGGCAGGCTGAGCGGCGACATCACTTGAGTC 8184
|||||
DB 1 TGCTACATCACTATATATCTAGCACTTTGGAGGCTGATGTGGCGGATCACTGAGGTC 60
OY 8185 AGGATTTGAGACACGAGCTGGCCACATGTTGAACCTGCTCCACATA 8233
|||||
DB 61 GGGAGTTGAGACACGAGCTGGCCACATGTTGAACCTGCTCCACATA 109

RESULT	12
LOCUS	B17434
DEFINITION	345K2.TVB CIT978SKA1 Homo sapiens genomic clone A-345K02, genomic survey sequence.
ACCESSION	B17434
VERSION	B17434.1
KEYWORDS	GSS.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 (bases 1 to 109)
TITLE	Adams,M.D., Kelley,J.M., Rounsley,S.R. and Venter,J.C.
JOURNAL	Use of a BAC End Sequence Database for Sequence-Ready Map Building Unpublished (1997)
COMMENT	Other_GSSs: 345K02..TP 345K02..TPB Contact: Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: mdadams@tigr.org Clones are available from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html Seq primer: T7 Class: BAC ends.
FEATURES	location/Qualifiers
source	1..109
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone="A-345K02"
	/clone_1lb="CIT978SKA1"
	/sex="Female"
	/cell_type="Fibroblast"
	/note="Vector: pBAC108L; site_1: HindIII; site_2: HindIII;
	Caltech Human BAC Library A1"
BASE COUNT	24 a 30 c 31 g 24 t
ORIGIN	
Query Match	0.5%; Score 86.6; DB 84; Length 109;
Best Local Similarity	87.2%; Pred. No. 0.34; Matches
Matches	95; Conservative 0; Mismatches 14; Indels 0; Gaps 0.
OY	11325 TAGTAGAGATGGGTTTCACCATGTGTGGCCAGATGATCTCGAACACTGCAGCTCAAGTG 11384
Db	109 TAGTGTGAAGCGGGGTTTCACCATGTGTGGCCAGCTGTGCTCGAACTCCGCACTCATGAGTG 50
OY	11385 ATCCGCCACCGCTCATCTCCCAAATGCGTCGGGATTACAGGTGTGAGCCA 11433
Db	49 ATCCGCCACCATCAGCTCCCAAAGTGTGATGATTATAGTATGAGCCA 1
RESULT	13
LOCUS	AA244245
DEFINITION	AA244245 110 bp mRNA EST 20-AUG-1997
ACCESSION	nc07a04.s1 NCI-CGAP.Pr1 Homo sapiens cDNA clone IMAGE:1007406
VERSION	similar to contains Alu repetitive element,, mRNA sequence.
KEYWORDS	AA244245
SOURCE	AA244245.1 GI:1875104
ORGANISM	human.
REFERENCE	Homo sapiens
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
	National Cancer Institute (CGAP),
	Tumor Gene Index

JOURNAL
COMMENT

Unpublished (1997)
On Jan 24, 1995 this sequence version replaced g1:634306.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuéqui,
M.D., Michael Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Kitzman, Ph.D.
cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
www.bio.llnl.gov/bdrrp/image/image.html

Seq primer: -41m13 fwd, ET from Amersham
High quality sequence stop: 90.
Location/Qualifiers
1. 110
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1007406"
/clone_id="NCI_CGAP_Prl"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: pAMP10; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected, histologically normal prostate epithelial cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into pAMP10 by the UDE-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."

BASE COUNT 17 a 26 c 28 g 38 t 1 others
ORIGIN

Query Match 0.5%; Score 86.6; DB 30; Length 110;
Best Local Similarity 86.4%; Pred. No. 0.34; Indels 0; Gaps 0;
Matches 95; Conservative 0; Mismatches 15;

Qy 13412 TTTTGTGGAGATGCATCTTCATCCACCGAGTGAGACAGTGCATGATCT 13471
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 TTTTGTGGAGATGCATCTTCATCCTGTTGCCAGCTGAGTGCAGTGCAGANTCT 60
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 13472 TGCGTCATTGCACCTCGCTCCTCGGTTCAAGTAGATTCTCGATCTCA 13521
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 TGGCTACTGCACACCTCGCTCCTCGGTTCAAGTAGATTCTCTCCYCA 110
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 14
A0003188 LOCUS 110 bp DNA GSS 14-APR-1999
DEFINITION RPEC11-1D10.TPN RPCI-11 Homo sapiens genomic clone RPEC1-11-1D10,
genomic survey sequence.
ACCESSION A0003188
VERSION A0003188.1 GI:3030392
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 110)
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C.C., de Jong,P. and
Venter,J.C.
Use of BAC End Sequences for Sequence-Ready Map Building (1998)
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department Of Eukaryotic Genomics

The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208

Email: mdadams@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.

FEATURES
source Location/Qualifiers

1..110
/organism="Homo sapiens"
/db_xref="GDB:7500081"
/db_xref="taxon:9606"
/clone="RPCI-11-1D10"
/clone_11b="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
RPC11 Human Male BAC Library"
BASE COUNT 22 a 27 c 26 g 35 t
ORIGIN

Query Match 0.5%; Score 86.6; DB 94; Length 110;
Best Local Similarity 87.2%; Pred. No. 0.34; Indels 0; Gaps 0;
Matches 95; Conservative 0; Mismatches 14;

QY 12868 TTTTGTATTATTAGATGAGATGAGTTGCGGTGTAGCCAGATGCTCGATCTCC 12927
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2 TTTTGTATTATTAGATGAGATGAGTTTACCATGTTGGCAGATGCTCCGATCTCT 61
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 12928 TGACCTCGATCGACCGGCTCGGCTCCCAAGTGTGGATTACAG 12976
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 62 TGACCTCATGATCCACCTGCGCGAGCTCCCAAGTGTGGATTACAG 110
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 15
AA565533 107 bp mRNA EST 08-SEP-1997
LOCUS nk42b1.s1 NCI-CGAP_GC2 Homo sapiens cDNA clone IMAGE:1016157 3'
DEFINITION similar to contains Alu repetitive element; mRNA sequence.
ACCESSION AA565533
VERSION AA565533.1 GI:2337172
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 107)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Sep 12, 1996 this sequence version replaced gi:1393355.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Stratagene, Inc., David B. Krizman,
Ph.D.
cDNA Library Arraying: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNI at:
www.bio.llnl.gov/dbrrp/image/image.html
Insert Length: 1661 Std Error: 0.00

Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 87.
Location/Qualifiers

FEATURES
source Location/Qualifiers

1..107
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1016157"
/clone_11b="NCI-CGAP_GC2"
/tissue_type="germ cell tumor"
/lab_host="SOLR (kanamycin resistant)"
/note="vector: Bluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dt. Bulk
germ cell tumor. 5' adaptor sequence: 5' GAATTCGGCAG 3'
3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'
Average insert size: 1.2 kb."
BASE COUNT 22 a 34 c 26 g 25 t
ORIGIN

Query Match 0.5%; Score 86.2; DB 35; Length 107;
Best Local Similarity 87.9%; Pred. No. 0.38; Indels 0; Gaps 0;
Matches 94; Conservative 0; Mismatches 13;

QY 4311 TGGTCCACCTGTATATCCACCTACTCAGAGGCTGAAGCGGAGATCACTGAACT 4370
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 107 TGGTGTGCTCCTGTATATCCACCTACTCAGAGGCTGAGGACGACGAACTGAACCT 48
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 4371 GGGAGGCGAGATGTCATGAGCTGATGCATGCCACGCTCCAG 4417
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 47 GGGAGGCGAGAGCTTGGCATGAGCTGAGATTGAGCCACTGCTCCAG 1
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Search completed: June 17, 2000, 02:34:49
Job time: 223056 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2000, 10:50:11 ; Search time 593.69 Seconds
(without alignments)
3941.217 Million cell updates/sec

Title: US-08-852-495C-1_COPY_196000_214000
Perfect score: 18001
Sequence: 1 GATAGGCTGACTCTTAACCA.....CCATTCAGCTTCTCACTCTT 18001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/5C_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/5D_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/6_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/PCtUS_COMB.seq:*
7: /cgn2_6/ptodata/1/ina/Backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	78.4	0.4	105	4	US-08-481-658B-65
2	78.4	0.4	105	4	US-08-477-504A-65
3	78.4	0.4	105	4	US-08-486-756A-65
4	78.4	0.4	105	4	US-08-485-862B-65
5	78.4	0.4	105	5	US-08-787-739-65
6	72.8	0.4	105	4	US-08-481-658B-65
7	72.8	0.4	105	4	US-08-477-504A-65
8	72.8	0.4	105	4	US-08-486-756A-65
9	72.8	0.4	105	4	US-08-485-862B-65
10	72.8	0.4	105	5	US-08-787-739-65
11	63.8	0.4	84	3	US-08-454-557C-91
12	63.8	0.4	84	4	US-08-340-426D-91
13	63.8	0.4	84	4	US-08-450-673C-91
14	63.8	0.4	84	6	PCT-US95-17111A-91
15	60.4	0.3	78	4	US-08-454-557C-70
16	60.4	0.3	78	4	US-08-340-426D-70
17	60.4	0.3	78	4	US-08-450-673C-70
18	60.4	0.3	78	6	PCT-US95-17111A-70
19	59.2	0.3	85	3	US-08-454-557C-92
20	59.2	0.3	85	4	US-08-340-426D-92
21	59.2	0.3	85	6	PCT-US95-17111A-92
22	59.2	0.3	85	6	PCT-US95-17111A-92
23	58.2	0.3	84	3	US-08-454-557C-91
24	58.2	0.3	84	4	US-08-340-426D-91
25	58.2	0.3	84	4	US-08-450-673C-91
26	58.2	0.3	84	6	PCT-US95-17111A-91
27	57.8	0.3	78	3	US-08-454-557C-70

C 28	57.8	0.3	78	4	US-08-340-426D-70	Sequence 70, Appl
C 29	57.8	0.3	78	4	US-08-450-673C-70	Sequence 70, Appl
C 30	57.8	0.3	78	6	PCT-US95-17111A-70	Sequence 70, Appl
C 31	57.4	0.3	85	3	US-08-332-768A-44	Sequence 44, Appl
C 32	55.4	0.3	76	3	US-08-454-557C-69	Sequence 69, Appl
C 33	55.4	0.3	76	4	US-08-340-426D-69	Sequence 69, Appl
C 34	55.4	0.3	76	4	US-08-450-673C-69	Sequence 69, Appl
C 35	55.4	0.3	76	6	PCT-US95-17111A-69	Sequence 69, Appl
C 36	53.6	0.3	60	3	US-08-454-557C-60	Sequence 60, Appl
C 37	53.6	0.3	60	4	US-08-340-426D-60	Sequence 60, Appl
C 38	53.6	0.3	60	4	US-08-450-673C-60	Sequence 60, Appl
C 39	53.6	0.3	60	6	PCT-US95-17111A-60	Sequence 60, Appl
C 40	52.6	0.3	60	3	US-08-454-557C-57	Sequence 57, Appl
C 41	52.6	0.3	60	4	US-08-340-426D-57	Sequence 57, Appl
C 42	52.6	0.3	60	4	US-08-450-673C-57	Sequence 57, Appl
C 43	52.6	0.3	60	6	PCT-US95-17111A-57	Sequence 57, Appl
C 44	52.2	0.3	98	1	US-08-088-658-42	Sequence 42, Appl
C 45	52.2	0.3	98	4	US-08-471-907A-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-08-481-658B-65
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08481,658B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3E
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-481-658B-65

Query Match 0.4%; Score 78.4; DB 4; Length 105;
Best Local Similarity 84.6%; Pred. No. 7e-08;

Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OY 12868 TTTTGTATTTTACTAGAGATGAGTTTCGCCGTGTTAGCCAGAGATGCTCGATCTCC 12927

DB 2 TTTTACATCTTTAGTAGAGACAGGAGTTTCACCATATTGGCCAGGCTGCTCAAACTCC 61

OY 12928 TGACCTCGATCCACCGGCTCGGCTCCCAAGTGTGGGAT 12971

DB 62 TGACCTGTGATCCACAGGCTCGGCTCTCCAAAGTGTGGGAT 105

RESULT 2

US-08-477-504A-65

; Sequence 65, Application US/08477504A

; Patent No. 5972353

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/477,504A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3D

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; US-08-477-504A-65

Query Match 0.4%; Score 78.4; DB 4; Length 105;

Best Local Similarity 84.6%; Pred. No. 7e-08;

Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OY 12868 TTTTGTATTTTACTAGAGATGAGTTTCGCCGTGTTAGCCAGAGATGCTCGATCTCC 12927

DB 2 TTTTACATCTTTAGTAGAGACAGGAGTTTCACCATATTGGCCAGGCTGCTCAAACTCC 61

OY 12928 TGACCTCGATCCACCGGCTCGGCTCCCAAGTGTGGGAT 12971

DB 62 TGACCTGTGATCCACAGGCTCGGCTCTCCAAAGTGTGGGAT 105

RESULT 3

US-08-486-756A-65

; Sequence 65, Application US/08486756A

; Patent No. 5981711

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,756A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; US-08-486-756A-65

Query Match 0.4%; Score 78.4; DB 4; Length 105;

Best Local Similarity 84.6%; Pred. No. 7e-08;

Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OY 12868 TTTTGTATTTTACTAGAGATGAGTTTCGCCGTGTTAGCCAGAGATGCTCGATCTCC 12927

DB 2 TTTTACATCTTTAGTAGAGACAGGAGTTTCACCATATTGGCCAGGCTGCTCAAACTCC 61

OY 12928 TGACCTCGATCCACCGGCTCGGCTCCCAAGTGTGGGAT 12971

DB 62 TGACCTGTGATCCACAGGCTCGGCTCTCCAAAGTGTGGGAT 105

RESULT 4

US-08-485-862B-65

; Sequence 65, Application US/08485862B

; Patent No. 5989838

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

```

; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,862B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
;
; US-08-485-862B-65
;
Query Match
Best Local Similarity 84.4%; Score 78.4; DB 4; Length 105;
Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
;
QY 12868 TTTTGTATTATTAGTAGAGATGGAGTTTCGCCGTGTAGCCAGAGATGCTCGATCTCC 12927
;
; ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; DB 2 TTTTACATCTTAGTAGAGACAGAGGTTTCACCATATTGGCCAGGCTCGCTCAAACTCC 61
;
QY 12928 TGACCTCGTATCCACCGCGCTCGGCTCCCAAGTCTGGGAT 12971
;
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; DB 62 TGACCTGTGATCCACGAGCTCGGCTCCCAAGTCTGGGAT 105
;
;
RESULT 5
; US-08-787-739-65
; Sequence 65, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
;

```

```

; APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
;
; US-08-787-739-65
;
Query Match
Best Local Similarity 84.6%; Score 78.4; DB 5; Length 105;
Matches 88; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
;
QY 12868 TTTTGTATTATTAGTAGAGATGGAGTTTCGCCGTGTAGCCAGAGATGCTCGATCTCC 12927
;
; ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; DB 2 TTTTACATCTTAGTAGAGACAGAGGTTTCACCATATTGGCCAGGCTCGCTCAAACTCC 61
;
QY 12928 TGACCTCGTATCCACCGCGCTCGGCTCCCAAGTCTGGGAT 12971
;
; ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
; DB 62 TGACCTGTGATCCACGAGCTCGGCTCCCAAGTCTGGGAT 105
;
;
RESULT 6
; US-08-481-658B-65/C
; Sequence 65, Application US/08481658B
; Patent No. 5955075
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 6 Mariposa Court
; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
;

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,658B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-481-658B-65

Query Match 0.4%; Score 72.8; DB 4; Length 105;
Best Local Similarity 86.8%; Pred. No. 1e-06;
Matches 92; Conservative 0; Mismatches 12; Indels 2; Gaps 1;

QY 8140 ATCCGACACTTGGCAGGCTGAGCGGCAGATCCTGAGCTCGAGCTTGAGACCA 8199
DB 105 ATCCGACACTTGGCAGGCTGAGCGGCAGATCCTGAGCTTGAGACCA 48
QY 8200 GCCTGGCAACATGTGTAACCCCTGCTCCACTAAATAACAAAA 8245
DB 47 GCCTGGCAACATGTGTAACCCCTGCTCCACTAAATAAGATGTAAAA 2

RESULT 7
US-08-477-504A-65/c
Sequence 65, Application US/08477504A
Patent No. 5972353
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,504A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:

NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-477-504A-65

Query Match 0.4%; Score 72.8; DB 4; Length 105;
Best Local Similarity 86.8%; Pred. No. 1e-06;
Matches 92; Conservative 0; Mismatches 12; Indels 2; Gaps 1;

QY 8140 ATCCGACACTTGGCAGGCTGAGCGGCAGATCCTGAGCTCGAGCTTGAGACCA 8199
DB 105 ATCCGACACTTGGCAGGCTGAGCGGCAGATCCTGAGCTTGAGACCA 48
QY 8200 GCCTGGCAACATGTGTAACCCCTGCTCCACTAAATAACAAAA 8245
DB 47 GCCTGGCAACATGTGTAACCCCTGCTCCACTAAATAAGATGTAAAA 2

RESULT 8
US-08-486-756A-65/c
Sequence 65, Application US/08486756A
Patent No. 5981711
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,756A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-486-756A-65

Query Match 0.4%; Score 72.8; DB 4; Length 105;
Best Local Similarity 86.8%; Pred. No. 1e-06;
Matches 92; Conservative 0; Mismatches 12; Indels 2; Gaps 1;

OY 8140 ATCCGACACTTTGGCAGCTGAGCGGCGAGATCATCTTGAGTCAGAGTTGAGACCA 8199
|||||
DB 105 ATCCGACACTTTGGGAGCGCGAGCTGCTGATCAC--AAGTCAGAGTTGAGAGCA 48
|||||

OY 8200 GCCTGGCCAACTGTTGTAACCCCTGCTCTCACTAAATAACAAAA 8245
|||||
DB 47 GCCTGGCCAACTGTTGTAACCCCTGCTCTCACTAAAGATGTAAAAA 2
|||||

RESULT 9

US-08-485-862B-65/C
Sequence 65, Application US/08485862B
Patent No. 5989838
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
CITY: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,862B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/477,504
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3D
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-485-862B-65

Query Match 0.4%; Score 72.8; DB 4; Length 105;
Best Local Similarity 86.8%; Pred. No. 1e-06;
Matches 92; Conservative 0; Mismatches 12; Indels 2; Gaps 1;

OY 8140 ATCCGACACTTTGGCAGCTGAGCGGCGAGATCATCTTGAGTCAGAGTTGAGACCA 8199
|||||
DB 105 ATCCGACACTTTGGGAGCGCGAGCTGCTGATCAC--AAGTCAGAGTTGAGAGCA 48
|||||

OY 8200 GCCTGGCCAACTGTTGTAACCCCTGCTCTCACTAAATAACAAAA 8245
|||||
DB 47 GCCTGGCCAACTGTTGTAACCCCTGCTCTCACTAAAGATGTAAAAA 2
|||||

RESULT 10

US-08-787-739-65/C
Sequence 65, Application US/08787739
Patent No. 6027887
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 96
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 369 Pine Street, Suite 610
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/787,739
FILING DATE: 24-JAN-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,049
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/486,756
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/481,658
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/485,862
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/487,077
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-981-2034
TELEFAX: 415-981-0332
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-787-739-65

Query Match 0.4%; Score 72.8; DB 5; Length 105;
Best Local Similarity 86.8%; Pred. No. 1e-06;
Matches 92; Conservative 0; Mismatches 12; Indels 2; Gaps 1;

Qy 8140 ATCCAGACCTTTGGCAGGCTGAGCGGACATCTGAGCTGAGGCTTTGAGACCA 8199

Db 105 ATCCAGACCTTTGGCAGGCTGAGCGGACATCTGAGCTGAGGCTTTGAGACCA 48

Qy 8200 GCCTGCGCAATGATGTAACCCCTGCTCCTACTAAATATACAAAA 8245

Db 47 GCCTGCGCAATGATGTAACCCCTGCTCCTACTAAATATATAAAA 2

RESULT 11
US-08-454-557C-91
; Sequence 91, Application US/08454557C

; Patent No. 5830670

; GENERAL INFORMATION:

; APPLICANT: de la Monte, Suzanne

; APPLICANT: Wands, Jack R.

; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection

; NUMBER OF SEQUENCES: 121

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

; STREET: 1100 New York Avenue, Suite 600

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20005-3934

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/454,557C

; FILING DATE: 30-MAY-1995

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Ludwig, Steven R.

; REGISTRATION NUMBER: 36,203

; REFERENCE/DOCKET NUMBER: 0609,3840003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 371-2600

; TELEFAX: (202) 371-2540

; INFORMATION FOR SEQ ID NO: 91:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 84 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: both

; TOPOLOGY: both

; US-08-454-557C-91

Query Match 0.4%; Score 63.8; DB 3; Length 84;
Best Local Similarity 85.5%; Pred. No. 6.8e-05;

Matches 71; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 12899 CCGGTGTACCGAGATGCTCTGATCTCTGACCTGTGATCCACCGCGCTCGCTCC 12958

Db 1 CCATGTTCTATCAGGCTGGTGTGACTCTGACCTGTGATCCGCGCTCGCTCC 60

Qy 12959 AAAGTCTGGGATTAACAGGCATG 12981

Db 61 AAAGTCTGGGATTAACAGCGTG 83

RESULT 12
US-08-340-426D-91
; Sequence 91, Application US/08340426D
; Patent No. 5948634

; GENERAL INFORMATION:

; APPLICANT: de la Monte, Suzanne

; APPLICANT: Wands, Jack R.

; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection

; NUMBER OF SEQUENCES: 121

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

; STREET: 1100 New York Avenue, Suite 600

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20005-3934

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/340,426D

; FILING DATE: 14-NOV-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Ludwig, Steven R.

; REGISTRATION NUMBER: 36,203

; REFERENCE/DOCKET NUMBER: 0609,3840002

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202) 371-2600

; TELEFAX: (202) 371-2540

; INFORMATION FOR SEQ ID NO: 91:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 84 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: both

; TOPOLOGY: both

; US-08-340-426D-91

Query Match 0.4%; Score 63.8; DB 4; Length 84;
Best Local Similarity 85.5%; Pred. No. 6.8e-05;

Matches 71; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 12899 CCGGTGTACCGAGATGCTCTGATCTCTGACCTGTGATCCACCGCGCTCGCTCC 12958

Db 1 CCATGTTCTATCAGGCTGGTGTGACTCTGACCTGTGATCCGCGCTCGCTCC 60

Qy 12959 AAAGTCTGGGATTAACAGGCATG 12981

Db 61 AAAGTCTGGGATTAACAGCGTG 83

RESULT 13
US-08-450-673C-91
; Sequence 91, Application US/08450673C

; Patent No. 5948888

; GENERAL INFORMATION:

; APPLICANT: de la Monte, Suzanne

; APPLICANT: Wands, Jack R.

; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection

; NUMBER OF SEQUENCES: 121

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

; STREET: 1100 New York Avenue, Suite 600

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20005-3934

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/450,673C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609,3840004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-450-673C-91

Query Match          0.4%; Score 63.8; DB 4; Length 84;
Best Local Similarity 85.5%; Pred. No. 6,8e-05;
Matches 71; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

OY 12899 CCGGTTCAGCAGATGCTCGATCTCCTGACCTGATCCACGCGCTCGCTCCC 12958
DB      1 CCAATGTCATCAGCGCTGGTGTGCACTCCTGACCTGCTGATCGCCGCTCAGCTCCC 60

OY 12959 AAGTCTGGGATTACAGCATG 12981
DB      61 AAGTCTGGGATTACAGCATG 83

RESULT 14
PCT-US95-17111A-91
; Sequence 91, Application PC/TUS9517111A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/17111A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,426
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609,3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
```

```

; TOPOLOGY: both
;
PCT-US95-17111A-91
```

```

Query Match          0.4%; Score 63.8; DB 6; Length 84;
Best Local Similarity 85.5%; Pred. No. 6,8e-05;
Matches 71; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
```

```

OY 12899 CCGGTTCAGCAGATGCTCGATCTCCTGACCTGATCCACGCGCTCGCTCCC 12958
DB      1 CCAATGTCATCAGCGCTGGTGTGCACTCCTGACCTGCTGATCGCCGCTCAGCTCCC 60

OY 12959 AAGTCTGGGATTACAGCATG 12981
DB      61 AAGTCTGGGATTACAGCATG 83
```

```

RESULT 15
US-08-454-557C-70
; Sequence 70, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/454,557C
; FILING DATE: 30-MAY-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609,3840003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 78 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
;
US-08-454-557C-70
```

```

Query Match          0.3%; Score 60.4; DB 3; Length 78;
Best Local Similarity 85.9%; Pred. No. 0.00034;
Matches 67; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
```

```

OY 13556 ACCATGCGTGGTAATTTTGTATTTTACAGATGGGTTTCACCATGTTGACGAG 13615
DB      1 AACAGCGCCAGCTAATTTATTTATTTAGTAGAGATGGGTTTCTCCAGTTTCATCAGG 60

OY 13616 CTGGTCTCAACATCTCTGA 13633
DB      61 CTGGTCTCAACATCTCTGA 78
```

Wed Jun 21 14:43:38 2000

us-08-852-495c-1_copy_196000_214000.rml

Page 8

Search completed: June 17, 2000, 10:50:33
Job time: 251763 sec

RESULT	1	2
LOCUS	HSIDLRLN2	108 bp DNA PRI 20-MAY-1992
DEFINITION	Human LDL-receptor gene Intron 14 fragment (normal gene).	
ACCESSION	X05250	
VERSION	X05250.1 GI:34337	
KEYWORDS	Alu repetitive sequence; low density lipoprotein receptor.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS	1 (bases 1 to 108) Williamson, R., Beisiegel, U., Dunning, A., Havinga, J.R., Horsthemke, B., and Humphries, S.	
TITLE	Unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia	
JOURNAL	Eur. J. Biochem. 164 (1), 77-81 (1987)	
MEDLINE	87161901	
COMMENT	See X05252 for deletion junction Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.	
FEATURES	Location/Qualifiers	
source	1..108 /organism="Homo sapiens" /db_xref="taxon:9606"	
intron	1..108 /note="intron XIV fragment"	
BASE COUNT	28 a 23 c 39 g 18 t	
ORIGIN		
Query Match	0.4%; Score 98.4; DB 10; Length 108;	
Best Local Similarity	94.4%; Pred. No.3.2e-06;	
Matches 102; Conservative 0; Mismatches 6; Indels 0; Gaps 0;		
Qy 14586	CTTGCGTCACGTCACACCTCTGCTCCTCGGTTCAGACAAATTCCTGCTCAGCCTTCCA	14645
Db 108	CTCGGCTCAGTCACACCTCTGCTCCTCGGTTCAGACAAATTCCTGCTCAGCCTCCGC	49
Qy 14646	AGTACTGCGGATTACAGCACCACCTACCACGCGCTGCTAATTTTGT	14693
Db 48	AGTAGCTGGGATTACAGCACCACCTACCACGCGCTGCTAATTTTGT	1
RESULT	2	
HUMALCE162/c		
LOCUS	HUMALCE162	107 bp ss-RNA PRI 15-APR-1994
DEFINITION	Human carcinoma cell-derived Alu RNA transcript, clone CE162.	
ACCESSION	M87924	
VERSION	M87924.1 GI:174871	
KEYWORDS	Alu repeat.	
SOURCE	Homo sapiens male embryo carcinoma cDNA to other RNA.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS	1 (bases 1 to 107) Sinnott, D., Richer, C., Deragon, J.-M. and Labuda, D.	
TITLE	Alu RNA transcripts in human embryonal carcinoma cells. Model of post-transcriptional selection of master sequences	
JOURNAL	J. Mol. Biol. (1992) in press	
FEATURES	Location/Qualifiers	
source	1..107 /organism="Homo sapiens" /db_xref="taxon:9606" /cell_line="Ntera2D1" /dev_stage="embryo" /sex="male" /tissue_type="carcinoma"	
BASE COUNT	28 a 30 c 35 g 14 t	
ORIGIN		

Query Match	0.4%;	Score 90.8;	DB 9;	Length 107;
Best Local Similarity	93.1%;	Pred. No. 4.9e-05;		
Matches	95;	Conservative	0;	Mismatches 7; Indels 0; Gaps 0;
OY	11779	TTTTTGAGACGGAGTCTCAGTCTGTGTACCCAGCGCTGAGTGCAGTGGCGATCTCGGCT	11838	
Db	106	TTTTTGAGACGGAGTCTCAGTCTGTGTGTACCCAGCGCTGAGTGCAGTGGCGATCTCGGCT	47	
OY	11839	CAGTGCAGACCTCCGCTCCCGGGTTCAAGTATTTCTCTGCC	11880	
Db	46	CAGTGCAGACCTCCGCTCCCGGGTTCAAGTATTTCTCTGCC	5	
RESULT	3			
LOCUS	HSIDLRLN2	108 bp	DNA	PRI
DEFINITION	Human LDL-receptor gene intron 14 fragment (normal gene).			20-MAY-1992
ACCESSION	X05250			
VERSION	X05250.1 GI:34337			
KEYWORDS	Alu repetitive sequence; low density lipoprotein receptor.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;			
AUTHORS	Primates; Catarrhini; Hominiidae; Homo.			
TITLE	1 (bases 1 to 108)			
JOURNAL	Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,			
MEDLINE	Williamson,R. and Humphries,S.			
COMMENT	Unequal crossing-over between two alu-repetitive DNA sequences in the low-density-lipoprotein-receptor gene. A possible mechanism for the defect in a patient with familial hypercholesterolaemia			
FEATURES	Eur. J. Biochem. 164 (1), 77-81 (1987)			
source	87161801			
	See X05253 for deletion junction			
	Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.			
	Location/Qualifiers			
	1..108			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	1..108			
	/note="intron XIV fragment"			
	28 a 23 c 39 g 18 t			
BASE COUNT				
ORIGIN				
	intron			
	Query Match	0.4%;	Score 90.4;	DB 10;
	Best Local Similarity	89.8%;	Pred. No. 5.6e-05;	
	Matches	97;	Conservative	0;
			Mismatches	11; Indels 0; Gaps 0;
OY	4976	ACAAAATTTAGCCAGAGTGTGTGTGGCAGGCACCTGTAAATCCACAGCTATTCAGAGGCTGAG	5035	
Db	1	ACAAAATTTAGCCAGAGCTGTGTGTGGCAGGCTGTAAATCCACAGCTATTCAGAGGCTGAG	60	
OY	5036	ACAGAGAGATGCTGTGAACCCAGAGAGTGGAGCTGCAATTTGAGCCAG	5083	
Db	61	GCAGAGAGATGCTGTGAACCCAGAGAGGAGAGGCTTGCAGTGAAGCCAG	108	
RESULT	4			
LOCUS	HSU67803/c	108 bp	RNA	PRI
DEFINITION	Human small cytoplasmic Alu transcript.			01-AUG-1997
ACCESSION	U67803			
VERSION	U67803.1 GI:2289917			
KEYWORDS	Alu.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;			
AUTHORS	Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
TITLE	1 (bases 1 to 108)			
JOURNAL	Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.			
	CDNAs derived from primary and small cytoplasmic Alu (scAlu)			
	transcripts			
JOURNAL	J. Mol. Biol. 271 (2), 222-234 (1997)			

MEDLINE 97415756
REFERENCE 2 (bases 1 to 108)
AUTHORS Shalh, T.H., Kim, J., Batzer, M.A. and Delinger, P.L.
TITLE Direct Submission
JOURNAL Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The Children's Hospital of Philadelphia, 104F Abramson Research Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA

FEATURES
source 1..108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="TSCALu2"
1..108
/note="scalu"
/rpt_family="Alu"
/rpt_type="dispersed"

BASE COUNT 23 a 39 c 30 g 16 t
ORIGIN

Query Match 0.4%; Score 89.6; DB 11; Length 108;
Best Local Similarity 95.8%; Pred. No. 7.5e-05;
Matches 92; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 10437 GTAGAGACGGGGTTTCACATGTTAGCCAGATGTCGATCTCCGACCTGATCC 10496
|||||
DB 97 GTAGAGACGGGGTTTCACCTGTTAGCCAGATGTCGATCTCCGACCTGATCC 38
|||||

QY 10497 GCCCAGCTGAGCCTCCCAAGTGTGGATTCAGG 10532
|||||
DB 37 GCCCGCTGGCTCCCAAGTGTGGATTCAGG 2
|||||

RESULT 5
HSLDLR1 108 bp DNA PRI 20-MAY-1992
LOCUS Human LDL-receptor mutated gene with intron 12 deletion junction.
DEFINITION X05249
ACCESSION X05249.1 GI:34335
VERSION X05249.1 GI:34335
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 108)
Horschmke, B., Beisiegel, U., Dunning, A., Havinga, J.R.,
Williamson, R. and Humphries, S.
Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia
Eur. J. Biochem. 164 (1), 77-81 (1987)

JOURNAL MEDLINE 87161901
COMMENT *source: hypercholesterol aemia
See X05248 for corresponding normal gene sequence
In the defective LDL-receptor gene the deletion occurred between two
alu-repetitive sequences, that are in the same direction, the
deletion eliminates exons 13 and 14 and changes the reading frame
of the resulting spliced mRNA.
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
Location/Qualifiers
1..108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="blood leukocytes from a patient with familial"

misc_feature 1..108
/note="deletion junction region intron 12/ intron 15"

BASE COUNT 20 a 40 c 20 g 28 t
ORIGIN

Query Match 0.4%; Score 87.8; DB 10; Length 108;
Best Local Similarity 88.8%; Pred. No. 0.00014;
Matches 95; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 14587 TTGGCTCAGTCAACCTCTGCTCTGTTCAAGCAATCTCTGCTCAGCTTCCAA 14646
|||||
DB 2 TCGCTTCACCAACACTCTCTCTCTGTTCAACCAATTTCTCTGCTCAGCTTCCGA 61
|||||

QY 14647 GTAGCTGGATTACAGGACCACTACCAAGCTGCTAAATTTTGT 14693
|||||
DB 62 GTAGCTGGATTACAGGACCACTGCAACGCGCTAAATTTTGT 108
|||||

RESULT 6
HSLDLR2/c 108 bp DNA PRI 20-MAY-1992
LOCUS Human LDL-receptor mutated gene with intron 14 deletion junction.
DEFINITION X05251
ACCESSION X05251.1 GI:34336
VERSION X05251.1 GI:34336
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 108)
Horschmke, B., Beisiegel, U., Dunning, A., Havinga, J.R.,
Williamson, R. and Humphries, S.
Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia
Eur. J. Biochem. 164 (1), 77-81 (1987)

JOURNAL MEDLINE 87161901
COMMENT *source: hypercholesterol aemia
See X05250 for corresponding normal gene sequence
In the defective LDL-receptor gene the deletion occurred between two
alu-repetitive sequences, that are in the same direction, the
deletion eliminates exons 13 and 14 and changes the reading frame
of the resulting spliced mRNA.
Data kindly reviewed (07-DEC-1987) by HUMPHRIES S.
Location/Qualifiers
1..108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="blood leukocytes from a patient with familial"

INTRON
/note="intron XIV fragment"

BASE COUNT 28 a 20 c 40 g 20 t
ORIGIN

Query Match 0.4%; Score 87.8; DB 10; Length 108;
Best Local Similarity 88.8%; Pred. No. 0.00014;
Matches 95; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 14587 TTGGCTCAGTCAACCTCTGCTCTGTTCAAGCAATCTCTGCTCAGCTTCCAA 14646
|||||
DB 107 TCGCTTCACCAACACTCTCTCTGTTCAACCAATTTCTCTGCTCAGCTTCCGA 48
|||||

QY 14647 GTAGCTGGATTACAGGACCACTACCAAGCTGCTAAATTTTGT 14693
|||||
DB 47 GTAGCTGGATTACAGGACCACTGCAACGCGCTAAATTTTGT 1
|||||

RESULT 7
HUMALCE221/c 103 bp ss-RNA PRI 15-APR-1994
LOCUS HUMALCE221 Human carcinoma cell-derived Alu RNA transcript, clone CE221.
DEFINITION M87896
ACCESSION M87896.1 GI:174874
VERSION M87896.1 GI:174874
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 103)
Smettt, D., Richer, C., Deragon, J.-M. and Labuda, D.

[illegible]

```
source
1.108
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="Blood leukocytes from a patient with familial"
Intron
1.108
/note="Intron XIV fragment"
BASE COUNT      28 a      20 c      40 g      20 t
ORIGIN

Query Match      0.4%; Score 81.4; DB 10; Length 108;
Best Local Similarity 85.0%; Pred. No. 0.0014;
Matches 91; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

QY 4976 ACACAAATTAAGCAGGTGTGGTGGCAGGACCTGTATATCCAGCTATTTCAGAGAGCTGAG 5035
DB 1 ACACAAATTAAGCAGGCGGTGGTGGCAGGTGTATATCCAGCTGTCTGGAGGCTGAG 60

QY 5036 ACAGAGATCGCTTGACCCAGAGAGTGGAGTTCATTGAGCCAA 5082
DB 61 GCAGGAAATGTTGTAACCCAGAGAGAGAGTGTGTGAGGCCA 107

RESULT 11
LOCUS      HSU67804      108 bp      RNA      PRI      01-AUG-1997
DEFINITION Human small cytoplasmic Alu transcript.
ACCESSION  U67804
VERSION     U67804.1 GI:2289918
KEYWORDS    Alu.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1 (bases 1 to 108)
AUTHORS     Shaikh,T.H., Roy,A.M., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE       cDNAs derived from primary and small cytoplasmic Alu (scalu)
            transcripts
JOURNAL     J. Mol. Biol. 271 (2), 222-234 (1997)
MEDLINE     97415756
REFERENCE   2 (bases 1 to 108)
AUTHORS     Shaikh,T.H., Kim,J., Batzer,M.A. and Deininger,P.L.
TITLE       Direct Submission
JOURNAL     Submitted (22-AUG-1996) Human Genetics and Molecular Biology, The
            Children's Hospital of Philadelphia, 1004F Abramson Research
            Center, 34th and Civic Center Blvd., Philadelphia, PA 19104, USA
FEATURES
     . source
       1.108
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /clone="Tscalu3"
       repeat_region
         1.108
         /note="scalu"
         /rpt_family="Alu"
         /rpt_type="dispersed"
BASE COUNT      26 a      38 c      26 g      18 t
ORIGIN

Query Match      0.4%; Score 81.6; DB 11; Length 108;
Best Local Similarity 90.6%; Pred. No. 0.0013;
Matches 87; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 10437 GTAGACGCGGGTTCACCATGTTAGCCAGATGATCTCGATCTCTGACCTGTGATCC 10496
DB 97 GGAAAGACGGGGTTCACCATGTTAGCCAGATGATCTCGATCTCTGACCTGTGATCC 38

QY 10497 GCCCAGCTGAGCCTCCCAAGTCTGGGATTACAGG 10532
DB 37 TCCCGCTTGGCCTTCAAGTCTGGGATTACAGG 2

RESULT 12
```

```
HUMALCE162
LOCUS      HUMALCE162      107 bp ss-RNA      PRI      15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE162.
ACCESSION  M87924
VERSION     M87924.1 GI:174871
KEYWORDS    Alu repeat.
SOURCE      Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1 (bases 1 to 107)
AUTHORS     Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE       Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
JOURNAL     J. Mol. Biol. (1992) In press
FEATURES
     . source
       1.107
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /cell_line="Ntera2D1"
       /dev_stage="embryo"
       /sex="male"
       /tissue_type="carcinoma"
BASE COUNT      28 a      30 c      35 g      14 t
ORIGIN

Query Match      0.4%; Score 79.6; DB 9; Length 107;
Best Local Similarity 86.3%; Pred. No. 0.0027;
Matches 88; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 5035 GACAGAGATCGCTTGAAACCCAGAGAGTGGAGTTCATTGAGCCAAATCTGCACCT 5094
DB 5 GCGAGAGATGCGCTGGAACCCGCGAGGCGGAGCTTGACGATGCGCGCCACT 64

QY 5095 GCACCTCAGCCTGGCGAGCAGACTAGACTCCGTTCAAAA 5136
DB 65 GCACCTCAGCCTGGCGAGCAGAGACTCCGTTCAAAA 106

RESULT 13
LOCUS      HUMALCE221      103 bp ss-RNA      PRI      15-APR-1994
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE221.
ACCESSION  M87896
VERSION     M87896.1 GI:174874
KEYWORDS    Alu repeat.
SOURCE      Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1 (bases 1 to 103)
AUTHORS     Sinnett,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE       Alu RNA transcripts in human embryonal carcinoma cells. Model of
            post-transcriptional selection of master sequences
JOURNAL     J. Mol. Biol. (1992) In press
FEATURES
     . source
       1.103
       /organism="Homo sapiens"
       /db_xref="taxon:9606"
       /cell_line="Ntera2D1"
       /dev_stage="embryo"
       /sex="male"
       /tissue_type="carcinoma"
BASE COUNT      25 a      27 c      33 g      18 t
ORIGIN

Query Match      0.4%; Score 78.8; DB 9; Length 103;
Best Local Similarity 87.8%; Pred. No. 0.0036;
Matches 86; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 5006 CCTGTATCCCACTATTTCAGAGGCTGAGACAGAGGAATCGTTAACCAGAGAGTGG 5065
```

|||||
Db 6 CCTGAATCCAGTACAGGAGAGGAGATCGCTTGACCCGAGGCGG 65
QY 5066 AGGTGCGATTGAGCAAGATGCGACCTGCACTCCAG 5103
|||||
Db 66 AGGTGCGATTGAGCGAGATGCGCTGCACTCCAG 103
|||||
RESULT 14
HUMALCE43/c 110 bp ss-RNA PRI 15-APR-1994
LOCUS
DEFINITION Human carcinoma cell-derived Alu RNA transcript, clone CE43.
ACCESSION M87900.1 GI:174876
VERSION
KEYWORDS Alu repeat.
SOURCE Homo sapiens male embryo carcinoma cDNA to other RNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 110)
AUTHORS Slimet,D., Richer,C., Deragon,J.-M. and Labuda,D.
TITLE Alu RNA transcripts in human embryonal carcinoma cells. Model of
post-transcriptional selection of master sequences
JOURNAL J. Mol. Biol. (1992) In press
FEATURES
Location/Qualifiers
1..110
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_line="NTera2D1"
/dev_stage="embryo"
/sex="male"
/tissue_type="carcinoma"
BASE COUNT 27 a 31 c 34 g 18 t
ORIGIN
Query Match 0.4%; Score 78.8; DB 9; Length 110;
Best Local Similarity 84.0%; Pred. No. 0.0036;
Matches 89; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 10448 GTTTCACCATGTAGCGAGATGCTGATCCCTGACCTGATCGCCCACTGAG 10507
|||||
Db 110 GTTTCGTCATGTAGCGAGATGCTGATCCCTGACCTGATCGCCCACTGAG 51
|||||
QY 10508 CCTCCCAAGTCTGCGATTACGATGAGCCACCGCGCGCC 10553
|||||
Db 50 CCTCCCAAGTCTGCGATTACGATGAGCCACCGCGCGCC 5
|||||
RESULT 15
HSLDL12 108 bp DNA PRI 20-MAY-1992
LOCUS
DEFINITION Human LDL-receptor gene intron 12 fragment (normal gene) LDL = low
density lipoprotein.
ACCESSION X05248
VERSION X05248.1 GI:34334
KEYWORDS Alu repetitive sequence; low density lipoprotein receptor;
repetitive sequence.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Horsthemke,B., Beisiegel,U., Dunning,A., Havinga,J.R.,
Williamson,R. and Humphries,S.
TITLE Unequal crossing-over between two alu-repetitive DNA sequences in
the low-density-lipoprotein-receptor gene. A possible mechanism for
the defect in a patient with familial hypercholesterolaemia
Eur. J. Biochem. 164 (1), 77-81 (1987)
JOURNAL 87161901
MEDLINE
COMMENT see X05249 for deletion junction
DATA Kindly reviewed (07-DEC-1987) by HUMPHRIES S.
FEATURES
Location/Qualifiers

source 1..108
/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature complement(<1..65)
Intron /note="Alu repeat"
1..108
/note="intron XII fragment"
BASE COUNT 21 a 38 c 20 g 29 t
ORIGIN
Query Match 0.3%; Score 76.8; DB 10; Length 108;
Best Local Similarity 83.7%; Pred. No. 0.0074;
Matches 87; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
QY 18268 TCGGCTCACAGCACTCCACCTCGCGGTTCAAGTATTCCTGCGCAGCGCTGGA 18327
|||||
Db 2 TCGGCTCACAGCACTCCACCTCGCGGTTCAAGTATTCCTGCGCAGCGCTGGA 61
|||||
QY 18328 GTAGCTGGGATTACAGACGCTGTGTACACACACACTGCTAATTTT 18371
|||||
Db 62 GTAGCTGGGATTACAGACGCTGTGTACACACACACTGCTAATTTT 105
|||||
Search completed: June 17, 2000, 17:34:18
Job time: 276942 sec


```

RESULT 2
T24892/c
ID T24892 standard; cDNA to mRNA; 100 BP.
AC T24892:
DT 05-NOV-1996 (first entry)
DE Human gene signature HOMGS06998.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
human; cloning; mapping; non-biased library; diagnosis; detection;
cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN MO9514772-A1.
PD 01-JUN-1995.
PE 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
for diagnosis of abnormal cell function, by preparing cDNA that
reflects relative abundance of corresp. mRNA in specific human
tissues
PS Claim 1: Page 1720; 2245bp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 100 BP; 28 A; 22 C; 25 G; 22 T;

Query Match 0.3%; Score 67.2; DB 1; Length 100;
Best Local Similarity 78.8%; Pred. No. 0.074;
Matches 78; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

OY 11773 TATTATTTTGGAGGAGTCTCACTGTCACCGAGCGTGGAGTCAGTGGCGGATC 11832
DB 100 TTTGTTGTTTCAACAGAGTGTCTGTCACCGAGCGAGTGCAANGTGCAATC 41
OY 11833 TCGGCTCACTGCAACCTCCGCCCTCCGGGTTCAAGTGAT 11871
DB 40 TCAGCTNATTGCAAAATTCTGCGCTCCAGGTTCAAGGAT 2

RESULT 3
T26828
ID T26828 standard; cDNA to mRNA; 108 BP.
AC T26828:
DT 14-NOV-1996 (first entry)
DE Human gene signature HOMGS09078.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
human; cloning; mapping; non-biased library; diagnosis; detection;
cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN MO9514772-A1.
PD 01-JUN-1995.
PE 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.

```

```

PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1: Page 2182; 2245bp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 108 BP; 18 A; 33 C; 23 G; 28 T;

Query Match 0.3%; Score 66; DB 1; Length 108;
Best Local Similarity 90.8%; Pred. No. 0.11;
Matches 69; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 1728 ATCTCTGACCTCGATGATCCGCCCTCGGCTCCCAAGTGTACAGGCGTG 1787
DB 2 ATCTCTGACCTGTATGCTGATCCGCCGCTGTCCTCCATATGCTGATTACAGGCATG 61
OY 1788 AGCCACCGCGCCCGCG 1803
DB 62 AGCCACCGCGCCCGCG 77

RESULT 4
T25009
ID T25009 standard; cDNA to mRNA; 108 BP.
AC T25009:
DT 07-NOV-1996 (first entry)
DE Human gene signature HOMGS07131.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
human; cloning; mapping; non-biased library; diagnosis; detection;
cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN MO9514772-A1.
PD 01-JUN-1995.
PE 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI: 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
for diagnosis of abnormal cell function, by preparing cDNA that
reflects relative abundance of corresp. mRNA in specific human
tissues
PS Claim 1: Page 1748; 2245bp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

```

```

50 Sequence 108 BP; 34 A; 31 C; 26 G; 15 T;

Query Match 0.3%; Score 63.4; DB 1; Length 108;
Best Local Similarity 73.8%; Pred. No. 0.26;
Matches 79; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

OY 5044 ATCGCTTAACCCAGGAGCTGGAGCTTCGATTTGAGCCCAAGATCATGTCACCTGCCTCCAG 5103
Db 2 ATCCGCTTAGGCCCAAGAGAGCCAGGCTCTCAGTGAACCCATGTGCACGACCATGATTTCCAG 61

OY 5104 CCTGGCGCAGAGTAAGACATCCGTTTCAAAAACAAAAACCCACAGA 5150
Db 62 CCTGAGTACAGACAGCAAGACCCTGTTGAAACAACACACACANCAA 108

RESULT 5
ID T25854/c
T25854 standard; cDNA to mRNA; 91 BP.
AC T25854;
DT 22-Oct-1996 (first entry)
DE Human gene signature H0MG508084.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
human; cloning; mapping; non-biased library; diagnosis; detection;
cell typing; abnormal cell function; ss.
OS Homo sapiens.
ON W0514772-A1.
PD 01-JUN-1995.
PE 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
RA (MATSU) MATSUBARA K.
PI (OKUBO) OKUBO K.
PI Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 1944; 2245pp; Japanese.
PS A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-126837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
CC Sequence 91 BP; 18 A; 22 C; 28 G; 18 T;

Query Match 0.3%; Score 62.6; DB 1; Length 91;
Best Local Similarity 79.8%; Pred. No. 0.34;
Matches 71; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

OY 11783 TGAGACGAGTCTCACTGTGTCACCCAGGCTGGATGCGAGTTCGCGCTCACT 11842
Db 90 TGAGACAGAGTCTCACTGTGTCACCCAGGCTGGATGCGAGTTCGCGCTCACT 31

OY 11843 GCAACCTCCGCTCCCGGGTTCAAGTGAT 11871
Db 30 TGAACCTCTGCTCTTAGGCTCAAGTGAT 2

RESULT 6
ID T25009/c
T25009 standard; cDNA to mRNA; 108 BP.

```

AC	T25009; 1996	(first entry)
DE	07-NOV-1996	
DI	Human gene signature HDWG507131.	
KM	Gene signature; messenger RNA; mRNA; relative abundance; frequency;	
KW	human; cloning; mapping; non-biased library; diagnosis; detection;	
KW	cell typing; abnormal cell function; ss.	
OS	Homo sapiens.	
PN	WO9514772-A1.	
PD	01-JUN-1995.	
PE	11-NOV-1994; J01916.	
PR	12-NOV-1993; JP-355504.	
PA	(MATS/) MATSUBARA K.	
PA	(OKUB/) OKUBO K.	
PI	Matsubara K, Okubo K;	
PT	WPI; 95-206931/27.	
PT	Identifying gene signatures in 3'-directed human cDNA library - e.g.	
PT	for diagnosis of abnormal cell function, by preparing cDNA that	
PT	reflects relative abundance of corresp. mRNA in specific human	
PT	tissues	
PS	Claim 1; Page 1748; 2245bp; Japanese.	
CC	A single-stranded DNA (or its complementary strand or the corresp.	
CC	double-stranded DNA) which comprises one of the 7837 "GS" sequences	
CC	given in T19001-T26837 and which is able to hybridise to part of	
CC	human genomic DNA. cDNA or mRNA is claimed. The GS (gene signature)	
CC	sequences were obtained from 3'-directed cDNA libraries prepared	
CC	from various human tissues; synthesis of cDNA was initiated from the	
CC	3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-	
CC	untranslated sequence is unique to a particular mRNA species, almost	
CC	all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library	
CC	is constructed so as to reflect accurately the relative abundance of	
CC	different mRNAs in the particular tissue from which it was derived.	
CC	The appearance frequency of a given GS in a cDNA library can be	
CC	determined (esp. using primers and probes derived from the GS	
CC	sequences) as a means of diagnosing abnormal cell function or for	
CC	recognising different cell types.	
SQ	Sequence 108 BP; 34 A; 31 C; 26 G; 15 T;	
Query Match 0.3%; Score 61.8; DB 1; Length 108;		
Best Local Similarity 72.9%; Pred. No. 0.44;		
Matches 78; Conservative 0; Mismatches 29; Indels 0; Gaps 0		
OY	10253 YTTTYYTTTTTTTTTTTAAGACAGATCTTACTGTGCGCCGGGCTGAAGTCAGTG	10312
DB	108 TTGTGTTGTGTGTTGTTTCAACAGGGCTTGCTGTCACTCAGGCGTGAAATNCAGTG	49
OY	10313 GTCGCATCTTACGTCACGCAAGCTTGGCTGCCGGGCTCATGCCAT	10359
DB	48 GGCTGACCATGGCTCACCTCAGCCTTGGGCTCATGGGCTCAGGCGCAT	2
RESULT 7		
ID	T24892	
AC	T24892 standard; cDNA to mRNA; 100 BP.	
DT	05-NOV-1996 (first entry)	
DE	Human gene signature HDWG506998.	
KM	Gene signature; messenger RNA; mRNA; relative abundance; frequency;	
KW	human; cloning; mapping; non-biased library; diagnosis; detection;	
KW	cell typing; abnormal cell function; ss.	
OS	Homo sapiens.	
PN	WO9514772-A1.	
PD	01-JUN-1995.	
PE	11-NOV-1994; J01916.	
PR	12-NOV-1993; JP-355504.	
PA	(MATS/) MATSUBARA K.	
PA	(OKUB/) OKUBO K.	
PI	Matsubara K, Okubo K;	
PT	WPI; 95-206931/27.	
PT	Identifying gene signatures in 3'-directed human cDNA library - e.g.	
PT	for diagnosis of abnormal cell function, by preparing cDNA that	
PT	reflects relative abundance of corresp. mRNA in specific human	
PT	tissues	

PS Claim 1; Page 1720; 2245bp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-726837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 100 BP; 28 A; 22 C; 25 G; 22 T;

Query Match 0.3%; Score 60.8; DB 1; Length 100;
Best Local Similarity 74.7%; Pred. No. 0.61;
Matches 74; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

OY 5044 ATCGCTTGAAACCCAGAGGTGAGTTCATTGAGCCAAATCAGTCCACTGCCTCCAG 5103
|||||
DB 2 ATCGCTTGAAACCTGCGAGCAGAAATTTGCAATTCAGATTCACACCTGCTCCAG 61

OY 5104 CCTGGGCGACAGAGTAAGCTCGTTTCAAAAACAAAA 5142
|||||
DB 62 CCTGGGCGACAGAGTAAGCTCGTTTGAACAAACAAA 100

RESULT 8

T26213
AC T26213; standard; cDNA to mRNA; 103 BP.
DT T26213; 13-NOV-1996 (first entry)
DE Human gene signature HUMGS08452.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-81.
PD 01-JUN-1995.
PR 11-NOV-1994; J01916.
PF 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUBO/) OKUBO K.
PA Matsubara K, Okubo K;
DR WPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 2029; 2245bp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-726837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 103 BP; 33 A; 21 C; 25 G; 23 T;

Query Match 0.3%; Score 60; DB 1; Length 103;
Best Local Similarity 75.0%; Pred. No. 0.79;
Matches 75; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

OY 5044 ATCGCTTGAAACCCAGAGGTGAGTTCATTGAGCCAAATCAGTCCACTGCCTCCAG 5103
|||||
DB 2 ATCGCTTGAGTCCAGAGCTTGGTTACAGTACGATGAGCGACCTGCCTCCAG 61

OY 5104 CCTGGGCGACAGAGTAAGCTCGTTTCAAAAACAAAA 5143
|||||
DB 62 CCTGGGCGACAGAGTAAGATGCTTTAAGAAAAAAA 101

RESULT 9

T26728
ID T26728; standard; cDNA to mRNA; 97 BP.

AC T26728; 22-OCT-1996 (first entry)
DE Human gene signature HUMGS08978.
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-81.
PD 01-JUN-1995.
PR 11-NOV-1994; J01916.
PF 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PA (OKUBO/) OKUBO K.
PA Matsubara K, Okubo K;
DR WPI; 95-206931/27.

PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
PS Claim 1; Page 2158; 2245bp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-726837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 97 BP; 19 A; 27 C; 20 G; 28 T;

Query Match 0.3%; Score 59.6; DB 1; Length 97;
Best Local Similarity 81.0%; Pred. No. 0.91;
Matches 68; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

OY 12004 GATCTACCCACCTGCGCTCCCAAAAGTCTGGGATTACAGCATAGCCATCGCCCGCAG 12063
|||||
DB 1 GATCTACCCACCTGCGCTCCCAAAAGTCTGGGATTACAGCATAGCCATCGCCCGCAG 60

OY 12064 CCGTCTGTCTTTTATTTATTA 12087
| | | | |
DB 61 NCTGTACTAAGTCTTTTATTTTAA 84

RESULT 10

T26213/c
ID T26213; standard; cDNA to mRNA; 103 BP.

AC T26213; 13-NOV-1996 (first entry)
DE Human gene signature HUMGS08452.

KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KM human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN MO9514772-AL.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PI (OKUBO/) OKUBO K.
PI Matsubara K.; Okubo K;
DR MPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
tissues
PS Claim 1; Page 2029; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA. cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 103 BP; 33 A; 21 C; 25 G; 23 T;

Query Match 0.3%; Score 59.4; DB 1; Length 103;
Best Local Similarity 74.3%; Pred. No. 0.97;
Matches 75; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

OY 10259 TTTTCTTTTATGACAGAGCTTACTGTGCGCCGGGCGAGTGCAGTGCCTGCA 10318
DB 102 TTTTCTTTTATGACAGAGCTTACTGTGCGCCGGGCGAGTGCAGTGCCTGCA 43
DB 10319 TCTTAGCTCACTGAAGCTTGGCTCCCGGGTCAATGCAT 10359
DB 42 TCATAGCTCACTGAACCAACCTCTGAGCTCAATGAT 2

RESULT 11
T20927
ID T20927 standard; cDNA to mRNA; 103 BP.
AC T20927;
DT 24-JUL-1996 (first entry)
DE Human gene signature HUMGS02180.
KM Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-AL.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PI (OKUBO/) OKUBO K.
PI Matsubara K.; Okubo K;
DR MPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
tissues
PS Claim 1; Page 758-759; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences

CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA. cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 103 BP; 22 A; 27 C; 21 G; 31 T;

Query Match 0.3%; Score 58; DB 1; Length 103;
Best Local Similarity 73.0%; Pred. No. 1.5;
Matches 73; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

OY 18304 GATTCTCTCCCTCAGCCCTCTGAGTACGTTACAGACGTGTACACACCTGG 18363
DB 1 GATCCCTCCACCTCCACCTCCAGTGTGCTACAGTGTGCCACATGTCAG 60
OY 18364 CTATTCTTGTATTTAGTAGAGACAGGGTTCCACCGTG 18403
DB 61 CTGATTTTGTATTTTATTTAGTAGGACAGTATTTCTCCATG 100

RESULT 12
T20373/C
ID T20373 standard; cDNA to mRNA; 107 BP.
AC T20373;
DT 19-JUL-1996 (first entry)
DE Human gene signature HUMGS01525.
KM Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
OS Homo sapiens.
PN WO9514772-AL.
PD 01-JUN-1995.
PF 11-NOV-1994; J01916.
PR 12-NOV-1993; JP-355504.
PA (MATS/) MATSUBARA K.
PI (OKUBO/) OKUBO K.
PI Matsubara K.; Okubo K;
DR MPI; 95-206931/27.
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
tissues
PS Claim 1; Page 623; 2245pp; Japanese.
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA. cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
SQ Sequence 107 BP; 26 A; 29 C; 17 G; 29 T;

Query Match 0.3%; Score 56.4; DB 1; Length 107;
Best Local Similarity 72.6%; Pred. No. 2.6;
Matches 69; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2000, 02:34:49 ; Search time 6489.89 Seconds
(without alignments)
13761.207 Million cell updates/sec

Title: US-08-852-495C-1_COPY_213000_235033
22034
Sequence: 1 TTGGAGCTGATTTCGTGCAT.....TGTCTGTGTGTGTGTGTGTG 22034

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4857316 seqs, 2026611650 residues

Total number of hits satisfying chosen parameters: 156056

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

EST:*
1: em_est1:*
2: em_est2:*
3: em_est3:*
4: em_est4:*
5: em_est5:*
6: em_est6:*
7: em_est7:*
8: em_est8:*
9: em_est9:*
10: em_est10:*
11: em_est11:*
12: em_est12:*
13: em_est13:*
14: em_est14:*
15: em_est15:*
16: em_est16:*
17: em_est17:*
18: em_est18:*
19: em_est19:*
20: qb_est1:*
21: qb_est2:*
22: qb_est3:*
23: qb_est4:*
24: qb_est5:*
25: qb_est6:*
26: qb_est7:*
27: qb_est8:*
28: qb_est9:*
29: qb_est10:*
30: qb_est11:*
31: qb_est12:*
32: qb_est13:*
33: qb_est14:*
34: qb_est15:*
35: qb_est16:*
36: qb_est17:*
37: qb_est18:*
38: qb_est19:*
39: qb_est20:*
40: qb_est21:*
41: qb_est22:*
42: qb_est23:*
43: qb_est24:*
44: qb_est25:*

45: qb_est26:*
46: qb_est27:*
47: qb_est28:*
48: qb_est29:*
49: qb_est30:*
50: qb_est31:*
51: qb_est32:*
52: em_est20:*
53: em_est21:*
54: em_est22:*
55: em_est23:*
56: em_est24:*
57: em_est25:*
58: em_est26:*
59: qb_est33:*
60: qb_est34:*
61: qb_est35:*
62: qb_est36:*
63: qb_est37:*
64: qb_est38:*
65: em_est27:*
66: em_est28:*
67: em_est29:*
68: em_est30:*
69: qb_est39:*
70: qb_est40:*
71: qb_est41:*
72: qb_est42:*
73: qb_est43:*
74: qb_est44:*
75: em_est31:*
76: em_est32:*
77: em_est33:*
78: em_est34:*
79: qb_est45:*
80: qb_est46:*
81: qb_est47:*
82: qb_gss1:*
83: qb_gss2:*
84: qb_gss3:*
85: qb_gss4:*
86: em_gss1:*
87: em_gss2:*
88: em_gss3:*
89: em_gss4:*
90: qb_gss5:*
91: qb_gss6:*
92: qb_gss7:*
93: qb_gss8:*
94: qb_gss9:*
95: em_gss5:*
96: em_gss6:*
97: em_gss7:*
98: em_gss8:*
99: em_gss9:*
100: em_gss10:*
101: em_gss11:*
102: qb_gss10:*
103: qb_gss11:*
104: em_gss12:*
105: qb_gss12:*
106: qb_gss13:*
107: qb_gss14:*
108: qb_gss15:*
109: qb_gss16:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result Query

No.	Score	Match	Length	DB	ID	Description
1	96.4	0.4	106	37	AA703692	AA703692 ag81a10.r
2	94.8	0.4	109	30	AA243009	AA243009 zr25h02.s
3	92.4	0.4	103	38	AA807640	AA807640 nx08h05.s
4	92.6	0.4	108	84	B65160	B65160 CIT-HSP-201
5	92.4	0.4	110	39	AA897366	AA897366 am06h02.s
6	91.4	0.4	109	94	AQ028426	AQ028426 CIT-HSP-2
7	91	0.4	108	84	B65160	B65160 CIT-HSP-201
8	89.8	0.4	101	39	AA835205	AA835205 ak64h01.s
9	89.8	0.4	110	30	AA244245	AA244245 nc07a04.s
10	88.6	0.4	103	30	AA228795	AA228795 nc1e07.s
11	88.4	0.4	106	105	AQ264176	AQ264176 CITBI-E1-
12	87.8	0.4	103	84	B48914	B48914 RPCI11-4A12
13	87.8	0.4	107	35	AA565533	AA565533 nk42b11.s
14	86.6	0.4	103	108	AQ535244	AQ535244 RPCI11-3
15	86.8	0.4	103	108	AQ582186	AQ582186 RPCI11-4
16	87	0.4	103	108	AQ584425	AQ584425 RPCI11-4
17	86.8	0.4	106	38	AA812141	AA812141 cb48h02.s
18	86.8	0.4	106	94	AQ062963	AQ062963 CIT-HSP-2
19	86.4	0.4	105	30	AA218889	AA218889 zq15d04.s
20	86.4	0.4	107	39	AA828124	AA828124 od71a07.s
21	86.6	0.4	109	94	AQ028426	AQ028426 CIT-HSP-2
22	86.6	0.4	110	30	AA244245	AA244245 nc07a04.s
23	85.6	0.4	102	30	AA226656	AA226656 nc19f09.s
24	85.8	0.4	105	28	AA078003	AA078003 7H12D08 C
25	85.8	0.4	105	105	AQ282107	AQ282107 RPCI11-94
26	85.2	0.4	109	22	H11143	H11143 ym09c06.r1
27	84.8	0.4	104	29	AA129957	AA129957 zn86h04.r
28	84.8	0.4	104	29	AA129957	AA129957 zn86h04.r
29	84.8	0.4	107	33	AA385808	AA385808 EST99495
30	85	0.4	109	105	AQ265749	AQ265749 CITBI-E1-
31	84.6	0.4	107	103	AQ240182	AQ240182 CIT-HSP-2
32	83.8	0.4	108	94	AQ014433	AQ014433 CIT-HSP-2
33	83.4	0.4	101	39	AA835205	AA835205 ak64h01.s
34	83.4	0.4	102	94	AQ004934	AQ004934 CIT-HSP-2
35	83.6	0.4	109	24	N25299	N25299 yw52c09.s1
36	83.4	0.4	109	84	B17434	B17434 345K2.TVB C
37	83	0.4	101	33	AA81369	AA81369 EST94442
38	83.2	0.4	104	108	AQ544583	AQ544583 CITBI-E1-
39	83.2	0.4	107	24	N23686	N23686 yw46a02.s1
40	83.4	0.4	110	94	AQ003188	AQ003188 RPCI11-1D
41	82.6	0.4	106	108	AQ544957	AQ544957 CITBI-E1-
42	82.8	0.4	110	29	AA177157	AA177157 nc02g07.s
43	82.4	0.4	101	94	AQ076649	AQ076649 CIT-HSP-2
44	82	0.4	98	24	H67549	H67549 yue8f10.s1
45	82.2	0.4	104	105	AQ268072	AQ268072 RPCI11-73

ALIGNMENTS

RESULT 1
 LOCUS AA703692 106 bp mRNA EST 24-DEC-1997
 DEFINITION ag81a10.r1 Strategene hnt neuron (#937233) Homo sapiens cDNA clone IMAGE:1140858 5' similar to contains Alu repetitive element;; mRNA sequence.
 ACCESSION AA703692
 VERSION AA703692.1 GI:2713610
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 106)
 Hillier,L., Allen,M., Bowles,L., Dubuque,T., Giesel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wyllie,T., Waterston,R. and Wilson,R. Mashu-NCI human EST Project
 TITLE Unpublished (1997)
 JOURNAL On Sep 12, 1996 this sequence version replaced gi:1397630.

FEATURES

source
 1..106
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1140858"
 /clone_lib="Strategene hnt neuron (#937233)"
 /dev_stage="hnt neurons"
 /lab_host="SOLR (kanamycin resistant)"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dt.
 Differentiated, post mitotic hnt neurons. Average insert size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor sequence: 5' CTCGAGTCTTTTCTTTTCTTTT 3'."
 High quality sequence stop: 53.
 Location/Qualifiers

BASE COUNT

19 a 29 c 29 g 29 t

Query Match 0.4%; Score 96.4; DB 37; Length 106;
 Best Local Similarity 94.3%; Pred. No. 0.066; 6; Indels 0; Gaps 0;
 Matches 100; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 10431 TTTTGTGTGAGAGCGGGTTTACATGTTAGCAGAGATGTCGATCCTGACCTCG 10490
 |||||
 DB 1 TTTTGTGTGAGAGCGAGGTTTACCTGTTAGCCAGAGATGTCGATCCTGACCTCG 60
 |||||
 QY 10491 TGAATCCGCCACCTGAGCTCCCAAGTCTGGGATTAACAGTGTC 10536
 |||||
 DB 61 TGAATCCGCCCGCTCAAGTCTGGGATTAACAGTGTC 106
 |||||

RESULT 2
 LOCUS AA243009 109 bp mRNA EST 11-MAR-1998
 DEFINITION zr25h02.s1 Strategene hnt2 neuronal precursor 937230 Homo sapiens cDNA clone IMAGE:664467 3' similar to contains Alu repetitive element; contains element LTR1 repetitive element;; mRNA sequence.
 ACCESSION AA243009
 VERSION AA243009.1 GI:1873869
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 109)
 Hillier,L., Allen,M., Bowles,L., Dubuque,T., Giesel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wyllie,T., Waterston,R. and Wilson,R. Mashu-NCI human EST Project
 TITLE Unpublished (1997)
 JOURNAL On Dec 3, 1996 this sequence version replaced gi:1126669.
 CONTACT: Wilson RK
 CONTACT: Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -28m13 rev1 ET from Amersham
 High quality sequence stop: 102.

FEATURES
source
Location/Qualifiers
1. 109
/organism="Homo sapiens"
/db_xref="GDB:5426481"
/db_xref="taxon:9606"
/clone_image="664467"
/clone_lib="Stratagene NT2 neuronal precursor 937230"
/tissue_type="neuroepithelial cells"
/dev_stage="Ntera-2 neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: brain; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. Uninduced, exponentially growing neuroepithelial
cells (Ntera-2/cl.D1). Average insert size: 1.0 kb;
Uni-ZAP XR Vector: -5' adaptor sequence: 5' GAATCGGACAGAG
3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3' "

BASE COUNT
19 a 30 c 30 g 30 t

ORIGIN

Query Match
Best Local Similarity 93.4%; Score 94.8; DB 30; Length 109;
Matches 99; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 10431 TTTTAGTAGAGACGGGTTTACCATGTTACCAAGATGTCGATCTCGACCTCG 10490
|||||
Db 4 TTTTAGTAGAGACGGGTTTACCATGTTACCAAGATGTCGATCTCGACCTCG 63
|||||

QY 10491 TGATCCGCCACCTGAGCTCCCAAGTGTGGGATTACAGGTGTG 10536
|||||
Db 64 TGATCCGCCACCTGAGCTCCCAAGTGTGGGATTACAGGTGTG 109
|||||

RESULT 3
LOCUS AA807640 103 bp mRNA EST 05-MAR-1998
DEFINITION nx08b05.s1 NCI-CGAP GC3 Homo sapiens CDNA clone IMAGE:1255473 3'
similar to contains Alu repetitive element; mRNA sequence.
ACCESSION AA807640
VERSION AA807640.1 GI:2877108
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 103)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT On Jan 19, 1998 this sequence version replaced gi:2151346.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/dbp/image/image.html

FEATURES
source
Insert Length: 774 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 87.
Location/Qualifiers
1. 103
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="1255473"
/clone_lib="NCI-CGAP_GC3"
/tissue_type="pooled germ cell tumors"

/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker. 1st strand cDNA was prepared from 3 pooled
germ cell tumors, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is not normalized. Library was
constructed by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT
19 a 27 c 30 g 27 t

ORIGIN

Query Match
Best Local Similarity 94.1%; Score 92.4; DB 38; Length 103;
Matches 96; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 10436 AGTAGAGACGGGTTTACCATGTTAGCCAGATGTCGATCTCGACCTGATGC 10495
|||||
Db 2 AGTAGAGATGGGGTTTACCGTGTAGCCAGATGTCGATCTCGACCTGATGC 61
|||||

QY 10496 CGCCACCTGAGCTTCCCAAGTGTGGGATTACAGGTGTGA 10537
|||||
Db 62 CGCTCAGCTCGGCTCCCAAGTGTGGGATTACAGGTGTGA 103
|||||

RESULT 4
LOCUS B65160 108 bp DNA GSS 21-JUN-1998
DEFINITION CIT-HSP-2017G2.TFB CIT-HSP Homo sapiens genomic clone 2017G2,
genomic survey sequence.
ACCESSION B65160
VERSION B65160.1 GI:2639138
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 108)
AUTHORS Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wilde,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building
JOURNAL Unpublished (1997)
COMMENT Other-GSS: CIT-HSP-2017G2.TFB
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamad@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html
Seq primer: M13 Reverse
Class: BAC ends.
Location/Qualifiers
1. 108
/organism="Homo sapiens"
/db_xref="GDB:7043860"
/db_xref="taxon:9606"
/clone_image="2017G2"
/clone_lib="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pReloBAC11; Site_1: HindIII; Site_2:
HindIII"

BASE COUNT
26 a 27 c 34 g 21 t

ORIGIN

Query Match 0.4%; Score 92.6; DB 84; Length 108;
Best Local Similarity 91.6%; Pred. No. 0.17;
Matches 98; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4994 TGGTGGAGGCGACCTGATCCAGTATCCAGAGGCTGAGAGGAGAAATCGTTGAA 5053
||||| ||| ||||||||||| ||||||||||| ||||||||||| |||||||
DB 1 TGTGTGACAGCCCGCTGTATATCCAGTACTGAGAGGCTGAGAGAAATCGTTGAA 60
||||| ||| ||||||||||| ||||||||||| ||||||||||| |||||||

QY 5054 CCCAGAGGTGGAGTTCATTGAGCCAGATCATGCTGCTGCTC 5100
||| ||||||||||| ||||||||||| ||||||||||| |||||||
DB 61 CCGGGAGGTGAGGTTCAGTGCAGGACAGATCATGCTGCTGCTC 107
||||| ||| ||||||||||| ||||||||||| ||||||||||| |||||||

RESULT 5
AA897366 110 bp mRNA EST 04-JAN-1999
LOCUS AA897366
DEFINITION am6h02.s1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:1466067 3' similar to contains Alu repetitive element; mRNA
sequence.
ACCESSION AA897366
VERSION AA897366
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 110)
AUTHORS NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
On Jan 19, 1998 this sequence version replaced g1:2150764.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (Info@image.llnl.gov) for further information.
Insert Length: 834 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 63.
Location/Qualifiers
1. 110
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1466067"
/clone_1lb="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NbHL19W, testis NHT, and B-cell
NCI CGAP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
1.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT 22 a 27 c 29 g 32 t
ORIGIN

Query Match 0.4%; Score 92.4; DB 39; Length 110;
Best Local Similarity 90.0%; Pred. No. 0.18;
Matches 99; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 10427 GTATTTTATAGAGACGCGGTTTCACCATGTAGCCAGATGTCGATCTCTGAC 10486
||||| ||| ||||||||||| ||||||||||| ||||||||||| |||||||
DB 1 GTATTTTATAGAGATGGGTTTCACCGTGTAAACAGATGTCGATCTCTGAGAC 60
||||| ||| ||||||||||| ||||||||||| ||||||||||| |||||||

QY 10487 CTCGCGATCCGCCACCTGAGCTCCCAAAAGTCGTGGATTACAGGTGTCG 10536
||||| ||| ||||||||||| ||||||||||| ||||||||||| |||||||

DB 61 CTCATGATCCGCCACCTGCGCTCCCAAGTCTGGATATAGCCGTG 110
||||| ||| ||||||||||| ||||||||||| ||||||||||| |||||||

RESULT 6
A0028426 109 bp DNA GSS 30-JUN-1998
LOCUS A0028426
DEFINITION CIT-HSP-2313615.TF CIT-HSP Homo sapiens genomic clone 2313615,
genomic survey sequence.
ACCESSION A0028426
VERSION A0028426
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 109)
AUTHORS Adams,M.D., Rounsley,S.D., Zhu,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdamas@ligr.org
Clones are available from Research Genetics (Info@resgen.com). BAC
end search page:
http://www.ligr.org/cdb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
Location/Qualifiers
1. 109
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2313615"
/clone_1lb="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"

BASE COUNT 19 a 36 c 25 g 29 t
ORIGIN

Query Match 0.4%; Score 91.4; DB 94; Length 109;
Best Local Similarity 89.9%; Pred. No. 0.23;
Matches 98; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 17725 TTTTGTGAGACGAGCTGCTCTCTGACCCAGCGTGGAGTGCAGGACCAATCTC 17784
||||| ||| ||||||||||| ||||||||||| ||||||||||| |||||||

DB 1 TTTGTTTCTGAGACGAGCTGCTCTCTGACCCAGCGTGGAGTGCAGGACCAATCTC 60
||||| ||| ||||||||||| ||||||||||| ||||||||||| |||||||

QY 17785 GCGTCACGACACCTCCGCGGCTCCAGCATTTCTCTGCTCA 17833
||||| ||| ||||||||||| ||||||||||| ||||||||||| |||||||

DB 61 AGCTACTGACACCTCCGCGGCTCCAGCATTTCTCTGCTCA 109
||||| ||| ||||||||||| ||||||||||| ||||||||||| |||||||

RESULT 7
B65160 108 bp DNA GSS 21-JUN-1998
LOCUS B65160
DEFINITION CIT-HSP-201762.TRB CIT-HSP Homo sapiens genomic clone 201762,
genomic survey sequence.
ACCESSION B65160
VERSION B65160.1 GI:2639138
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE	1 (bases 1 to 108)					
AUTHORS	Adams,M.D., Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.					
TITLE	Use of a random BAC End Sequence Database for Sequence-Ready Map Building					
JOURNAL	Unpublished (1997)					
COMMENT	Other_GSSS: CIT-HSP-2017G2.TFF Contact: Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel: 301 838 0200 Fax: 301 838 0208 Email: mda@deneligr.org Clones are available from Research Genetics (Info@resgen.com). BAC and search page: http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html Seq primer: M13 Reverse Class: BAC ends.					
FEATURES						
SOURCE	Location/Qualifiers 1..108 /organism="Homo sapiens" /db_xref="GDB:7043860" /db_xref="taxon:9606" /clone="2017G2" /clone_1lb="CIT-HSP" /sex="Male" /cell_type="Sperm" /note="Vector: pBeloBAC11; Site_1: HindIII; Site_2: HindIII"					
BASE COUNT	26 a 27 c 34 g 21 t					
ORIGIN						
Query Match	0.4%; Score 91; DB 84; Length 108;					
Best Local Similarity	90.7%; Pred. No. 0.26;					
Matches	97; Conservative 0; Mismatches 10; Indels 0; Gaps 0					
Oy	18250 GAGTCAGGGGATGATCTCGGCTCACACCAACCCTCGTGGTTTAAGTAGTTCT 18309 Db 107 GGTCGCAGGGGATGATCTTGCTGCTACTGTCCMACCTCACCCTCCGGGTTCAGAGATTCT 48 Oy 18310 CCTGCCCTACGCTCTCTGAGTACCTGGGATTACAGACGTGTGCACCA 13356 Db 47 CCTGCCCTACGCTCTCTGAGTACCTGGGATTACAGGCGCATGCCCACCA 1					
RESULT	8					
LOCUS	AA835205 101 bp mRNA EST 23-FEB-1998					
DEFINITION	AK64h01.s1 Barstead pancreas HPLR1 Homo sapiens cDNA clone IMAGE:1412689.3' similar to contains Alu repetitive element::contains element KER repetitive element;; mRNA sequence.					
ACCESSION	AA835205					
VERSION	AA835205.1 GI:2908933					
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 101) Hillier,L., Allen,M., Bowles,L., Dubuque,T., Getzel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Thieling,B., White,Y., Wyllie,T., Waterston,R. and Wilson,R. WashU-NCI human EST Project Unpublished (1997) On Nov 29, 1993 this sequence version replaced gi:536191. Contract: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800					
TITLE	JOURNAL COMMENT					

Fax: 314 286 1810
Email: estewatson.wustl.edu

This clone is available royalty-free through LNL; contact the IMAG Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m13 fwd, ET from Amersham.

FEATURES
Location/Qualifiers
1..101

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1412689"
/clone_1lb="Barstead pancreas HPLRB1"
/sex="female"
/dev_stage="adult, 34 years"
/lab_noset="DH10B"
/note="Organ: pancreas; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cdna was primed with a Not I - oligo(dT) prime [5].
TGTTACGATCTGCAGTCGGAGCGGCCGCCCCCTTTTTTTTTTTTTTTTTTTT
3]); double-stranded cdna was ligated to Eco RI adaptors [AAATCGCATCCTGT], digested with Not I and cloned into th
Not I and Eco RI sites of the modified pT73 vector.
Library constructed by Bob Barstead."

BASE COUNT
ORIGIN
14 a 36 c 27 g 24 t

Query Match
Best Local Similarity 93.1%; Score 89.8; DB 39; Length 101;
Matches 94; Conservative 0; Mismatches 7; Indels 0; Gaps 0

Qy 11783 TGAGACGAGCTCACTGTGTCACCAGCGCTGAGTGCAGTCCGACTCGGCTCACT 11842
|||||
Db 1 TGAACGAGATCTCACCTCTGCGCCAGCGCTGAGTGCAGTCCGCTCACT 60
|||||

Qy 11843 GCACACCTCCGCCCTCCGGGTTCAAGTAGATTCTTCCTGCCA 11883
|||||
Db 61 CCAAGCTCCGCTCCGGGTTCAAGCGCATTCCTGCCCTCA 101
|||||

RESULT 9
LOCUS AA244245 110 bp mRNA EST 20-AUG-1997
DEFINITION nc07804.s1 NCI-CGAP_prl Homo sapiens CDNA clone IMAGE:1007406
similar to contabins Alu repetitive element;; mRNA sequence.
ACCESSION AA244245
VERSION AA244245.1 GI:1875104
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Euarchia; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 110)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Jan 24, 1995 this sequence version replaced gi:634306.
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuagui,
M.D., Michael Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Kitzman, Ph.D.
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL, at:
www.dio.llnl.gov/db/rp/image/image.html

JOURNAL COMMENT
Seq primer: -4lm13 fwd, ET from Amersham
High quality sequence stop: 90.
Location/Qualifiers

FEATURES

source

1. 110
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1007406"
/clone_1lb="NCI_CGAP_Prl1"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: PAMP10; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected, histologically normal prostate epithelial cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into PAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."

BASE COUNT 17 a 26 c 28 g 38 t 1 others
ORIGIN

Query Match 0.4%; Score 89.8; DB 30; Length 110;
Best Local Similarity 88.2%; Pred. No. 0.35;
Matches 97; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 17724 TTTTGTGAGACGAGTCTGTTCTGTCACCCAGGCTGGAGTGCAGCACTCT 17783
|||||
DB 1 TTTTGTGAGATGAGATCTTGTGATCTTGTCCAGGCTGGAGTGCAGCACTCT 60
|||||

QY 17784 CGGCTCAGTCAAGCTCCGCTCCGGGTTCAAGCTATTCTCGCTCA 17833
|||||
DB 61 TGGCTCAGTCAAGCTCCGCTCCGGGTTCAAGATCTTCTGCTCA 110
|||||

RESULT 10
LOCUS AA228795 103 bp mRNA EST 20-AUG-1997
DEFINITION nc14e07.s1 NCI CGAP Prl Homo sapiens cDNA clone IMAGE:1008132
similar to contains Alu repetitive element; contains element MER28
repetitive element ;, mRNA sequence.
ACCESSION AA228795.1 GI:1851455
VERSION AA228795
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 103)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
On Sep 12, 1996 this sequence version replaced gi:1394473.
CONTACT: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuagui,
M.D., Michael Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LNM at:
www-bio.1nl.gov/dbtrp/image/image.html

JOURNAL
COMMENT

FEATURES
SOURCE

Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 81.
Location/Qualifiers
1. 103
/organism="Homo sapiens"
/db_xref="taxon:9606"

/clone="IMAGE:1008132"
/clone_1lb="NCI_CGAP_Prl1"
/sex="Male"
/dev_stage="45 years old"
/lab_host="DH10B"
/note="Vector: PAMP10; Site_1: NotI; Site_2: EcoRI; 1st strand cDNA was primed with oligo(dT)17 on 50 ng of DNase-treated, total cellular RNA obtained from 5,000-10,000 microdissected, histologically normal prostate epithelial cells. Double-stranded cDNA was ligated to EcoRI adaptors, 5 cycles of PCR applied to the cDNA with an adaptor-specific primer, and the resulting PCR product subcloned into PAMP10 by the UDG-cloning method (Life Technologies). Average insert size is 600 bp. NOTE: Not directionally cloned. This library was constructed by David Krizman."

BASE COUNT 14 a 30 c 25 g 34 t
ORIGIN

Query Match 0.4%; Score 88.6; DB 30; Length 103;
Best Local Similarity 91.3%; Pred. No. 0.48;
Matches 94; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1519 TTTTGTGAGACGAGTCTGTCACCTTTTCGCCAGCCGAGCTGCAGTGGCTAT 1578
|||||
DB 1 TTTTGTGAGATGAGTGTCTGTCACCTGTGCGCCAGGCTGGAGTGCAGTGGCAAT 60
|||||

QY 1579 CTGGCTCAGTCAAGCTCCGCTCCGGGTTCAAGCCATT 1621
|||||
DB 61 CTGGCTCAGTCAAGCTCCGCTCCGGGTTCAAGCCATT 103
|||||

RESULT 11
LOCUS AQ264176 106 bp DNA GSS 27-OCT-1998
DEFINITION CITR1-E1-2509A2.TF CITR1-E1 Homo sapiens genomic clone 2509A2,
genomic survey sequence.
ACCESSION AQ264176
VERSION AQ264176.1 GI:3792743
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 106)
Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,
Berry, K., Granger, D., Suh, E., Wille, C., Shizuya, R., Simon, W. and
Venter, J. C.
Use of a random human BAC End Sequence Database for Sequence-Ready
Map Building
Unpublished (1998)
Other GSSs: CITR1-E1-2509A2.TF
CONTACT: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: madams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html.
Seq primer: M13-21
Class: BAC ends.
Location/Qualifiers
1. 106
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2509A2"
/clone_1lb="CITR1-E1"
/sex="male"
/cell_type="sperm"


```

TITLE      Venter, J. C.
            Use of BAC End Sequences from Library RPC1-11 for Sequence-Ready
            Map Building
JOURNAL     Unpublished (1997)
COMMENT     On Feb 19, 1999 this sequence version replaced g1:4146076.
            Other GSSs: RPC1-11-451A15.TV
            Contact: Shaying Zhao, William Nierman, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbeetlgr.org
            Clones are derived from the human BAC library RPC1-11. For BAC
            library availability, please contact Pieter de Jong
            (pieterdejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genet cs (http://www.rtg.org/tldb/hungen/bac\_end\_search/bac\_end\_search.html).
            Seq primer: SP6
            Class: BAC ends.
FEATURES    location/Qualifiers
            source          1..103
                             /organism="Homo sapiens"
                             /db_xref="GDB:7672814"
                             /db_xref="taxon:9606"
                             /clone="RPC1-11-451A15"
                             /clone_id="RPC1-11"
                             /sex="Male"
                             /cell_type="Lymphocytes"
                             /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI;
                             RPC11 Human Male BAC Library"
BASE COUNT  19 a          36 c          25 g          22 t          1 others
ORIGIN
Query Match          0.43; Score 86.8; DB 108; Length 103;
Best Local Similarity 91.98; Pred. No. 0.76;
Matches 91; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY 17764 GGAGTGCAGTGGCACAACTCTGGGCTCACTGCAACCTCCCGGGTTCACGCTATTTC 17823
          |||||||
Db 1 GGAGTGCAGTGGCACAACTCTGGGCTCACTGCAACCTCTGCTCCCAAGATTCACAGGATTT 60
          |||||||
OY 17824 TCTGTGCTCAGCCCTCCCAAGTAGCTGGGAGACATACAGGCGC 17862
          |||||||
Db 61 TCTGTGCTCAGCCCTCCCAAGTAGCTGGGAGACATACAGGCGC 99
          |||||||

Search completed: June 17, 2000, 13:46:22
Job time: 263349 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 17, 2000, 10:50:33 ; Search time 286.25 Seconds
(without alignments)
10005.207 Million cell updates/sec

Title: US-08-852-495C-1_COPY_213000_235033
22034
Sequence: 1 TTCGAGTGAATTGGTCAT.....TGTGTGTGTGTGTGTGTGTG 22034

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 230463 seqs, 64992525 residues

Total number of hits satisfying chosen parameters: 374504

Minimum DB seq length: 10
Maximum DB seq length: 110

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/lna/5C.COMB.seq:*
4: /cgn2_6/ptodata/1/lna/5D.COMB.seq:*
5: /cgn2_6/ptodata/1/lna/6.COMB.seq:*
6: /cgn2_6/ptodata/1/lna/PCFUS.COMB.seq:*
7: /cgn2_6/ptodata/1/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76.2	0.3	105	4	US-08-481-658B-65 Sequence 65, Appl
2	76.2	0.3	105	4	US-08-477-504A-65 Sequence 65, Appl
3	76.2	0.3	105	4	US-08-486-756A-65 Sequence 65, Appl
4	76.2	0.3	105	4	US-08-485-862B-65 Sequence 65, Appl
5	76.2	0.3	105	5	US-08-787-739-65 Sequence 65, Appl
6	67	0.3	84	3	US-08-454-557C-91 Sequence 91, Appl
7	67	0.3	84	4	US-08-340-426D-91 Sequence 91, Appl
8	67	0.3	84	4	US-08-450-673C-91 Sequence 91, Appl
9	67	0.3	84	6	US-08-454-557C-91 Sequence 91, Appl
10	60.6	0.3	84	4	US-08-340-426D-91 Sequence 91, Appl
11	60.6	0.3	84	4	US-08-450-673C-91 Sequence 91, Appl
12	60.6	0.3	84	6	US-08-454-557C-91 Sequence 91, Appl
13	60.6	0.3	84	6	US-08-450-673C-91 Sequence 91, Appl
14	58.8	0.3	78	4	US-08-454-557C-70 Sequence 70, Appl
15	58.8	0.3	78	4	US-08-340-426D-70 Sequence 70, Appl
16	58.8	0.3	78	4	US-08-450-673C-70 Sequence 70, Appl
17	58.8	0.3	78	6	US-08-454-557C-70 Sequence 70, Appl
18	57.6	0.3	85	3	US-08-454-557C-92 Sequence 92, Appl
19	57.6	0.3	85	4	US-08-340-426D-92 Sequence 92, Appl
20	57.6	0.3	85	4	US-08-450-673C-92 Sequence 92, Appl
21	57.6	0.3	85	6	US-08-454-557C-92 Sequence 92, Appl
22	54.8	0.2	105	4	US-08-481-658B-65 Sequence 65, Appl
23	54.8	0.2	105	4	US-08-477-504A-65 Sequence 65, Appl
24	54.8	0.2	105	4	US-08-486-756A-65 Sequence 65, Appl
25	54.8	0.2	105	4	US-08-485-862B-65 Sequence 65, Appl
26	54.8	0.2	105	5	US-08-787-739-65 Sequence 65, Appl
27	54.4	0.2	106	3	US-08-332-766A-36 Sequence 36, Appl

28	54	0.2	76	3	US-08-454-557C-69	Sequence 69, Appl
29	54	0.2	76	4	US-08-340-426D-69	Sequence 69, Appl
30	54	0.2	76	4	US-08-450-673C-69	Sequence 69, Appl
31	54	0.2	76	6	US-08-454-557C-69	Sequence 69, Appl
32	53.6	0.2	60	3	US-08-454-557C-57	Sequence 57, Appl
33	53.6	0.2	60	4	US-08-340-426D-57	Sequence 57, Appl
34	53.6	0.2	60	4	US-08-450-673C-57	Sequence 57, Appl
35	53.6	0.2	60	6	US-08-454-557C-57	Sequence 57, Appl
36	50.4	0.2	60	5	US-08-340-426D-71	Sequence 71, Appl
37	50.4	0.2	60	6	US-08-454-557C-71	Sequence 71, Appl
38	50	0.2	94	5	US-08-750-064-6	Sequence 6, Appl
39	47.6	0.2	60	3	US-08-454-557C-60	Sequence 60, Appl
40	47.6	0.2	60	4	US-08-340-426D-60	Sequence 60, Appl
41	47.6	0.2	60	4	US-08-450-673C-60	Sequence 60, Appl
42	47.6	0.2	60	6	US-08-454-557C-60	Sequence 60, Appl
43	46.2	0.2	59	4	US-08-340-426D-65	Sequence 65, Appl
44	46.2	0.2	59	4	US-08-450-673C-65	Sequence 65, Appl
45	46.2	0.2	59	4	US-08-450-673C-65	Sequence 65, Appl

ALIGNMENTS

RESULT 1
US-08-481-658B-65
Sequence 65, Application US/08481658B
Patent No. 595075
GENERAL INFORMATION:
APPLICANT: Zavada, Jan
APPLICANT: Pastorekova, Silvia
APPLICANT: Pastorek, Jaromir
TITLE OF INVENTION: MN Gene and Protein
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESSEE: Leona L. Lauder
STREET: 6 Mariposa Court
City: Tiburon
STATE: California
COUNTRY: USA
ZIP: 94920
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,658B
FILING DATE: 07-JUN-1995
CLASSIFICATION: A24
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,190
FILING DATE: 15-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Lauder, Leona L.
REGISTRATION NUMBER: 30,863
REFERENCE/DOCKET NUMBER: D-0021.3E
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-435-2034
TELEFAX: 415-435-0727
INFORMATION FOR SEQ. ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 105 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-481-658B-65
Query Match 0.3%; Score 76.2; DB 4; Length 105;
Best Local Similarity 82.9%; Pred. No. 5.6e-07;

Matches 87; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 10422 TTTTGTATTTTATAGAGAGGGGTTTACCATGTTAGCAGAGATGCTCGATCTC 10481

Db 1 TTTTATACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTC 60

Qy 10482 CTGACCTGCTGATCCGCCACCTGAGCCTCCCAAAGTGTGGGAT 10526

Db 61 CTGACCTGTGATCCACACAGCCTCGGCTCCCAAAGTGTGGGAT 105

RESULT 2

US-08-477-504A-65

; Sequence 65, Application US/08477504A

; Patent No. 5972353

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentln Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/477,504A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3D

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; US-08-477-504A-65

Query Match 0.3%; Score 76.2; DB 4; Length 105;

Best Local Similarity 82.9%; Pred. No. 5.6e-07;

Matches 87; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 10422 TTTTGTATTTTATAGAGAGGGGTTTACCATGTTAGCAGAGATGCTCGATCTC 10481

Db 1 TTTTATACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTC 60

Qy 10482 CTGACCTGCTGATCCGCCACCTGAGCCTCCCAAAGTGTGGGAT 10526

Db 61 CTGACCTGTGATCCACACAGCCTCGGCTCCCAAAGTGTGGGAT 105

RESULT 3

US-08-486-756A-65

; Sequence 65, Application US/08486756A

; Patent No. 5981711

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

; CITY: Tiburon

; STATE: California

; COUNTRY: USA

; ZIP: 94920

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentln Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,756A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/260,190

; FILING DATE: 15-JUN-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Lauder, Leona L.

; REGISTRATION NUMBER: 30,863

; REFERENCE/DOCKET NUMBER: D-0021.3C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-435-2034

; TELEFAX: 415-435-0727

; INFORMATION FOR SEQ ID NO: 65:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 105 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; US-08-486-756A-65

Query Match 0.3%; Score 76.2; DB 4; Length 105;

Best Local Similarity 82.9%; Pred. No. 5.6e-07;

Matches 87; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 10422 TTTTGTATTTTATAGAGAGGGGTTTACCATGTTAGCAGAGATGCTCGATCTC 10481

Db 1 TTTTATACATCTTTAGTAGAGACAGGGTTTCACCATATTGGCCAGGCTGCTCTCAAACTC 60

Qy 10482 CTGACCTGCTGATCCGCCACCTGAGCCTCCCAAAGTGTGGGAT 10526

Db 61 CTGACCTGTGATCCACACAGCCTCGGCTCCCAAAGTGTGGGAT 105

RESULT 4

US-08-485-862B-65

; Sequence 65, Application US/08485862B

; Patent No. 5989838

; GENERAL INFORMATION:

; APPLICANT: Zavada, Jan

; APPLICANT: Pastorekova, Silvia

; APPLICANT: Pastorek, Jaromir

; TITLE OF INVENTION: MN Gene and Protein

; NUMBER OF SEQUENCES: 86

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Leona L. Lauder

; STREET: 6 Mariposa Court

```

; CITY: Tiburon
; STATE: California
; COUNTRY: USA
; ZIP: 94920
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,862B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER: US 08/260,190
; FILING DATE: 15-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.3D
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-435-2034
; TELEFAX: 415-435-0727
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
;
; US-08-485-862B-65

```

```

Query Match      0.3%; Score 76.2; DB 4; Length 105;
Best Local Similarity 82.9%; Pred. No. 5.6e-07;
Matches 87; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```

```

QY 10422 TTTTGTATTTTATAGAGAGGGGTTTCACCATGTTAGCCAGAGTGTCTGCATCTC 10481
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 TTTTTCATCTTTAGTAGAGAGGGTTTCACCATATTGGCCAGGCTGCTCCAACATC 60
QY 10482 CTGACCTCGTATCGGCCACCTGAGCCTCCCAAAGTCTGGGAT 10576
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 CTGACCTGTGTATCCACACAGCTCGGCTCCCAAAGTCTGGGAT 105

```

```

RESULT 5
US-08-787-739-65
; Sequence 65, Application US/08787739
; Patent No. 6027887
; GENERAL INFORMATION:
; APPLICANT: Zavada, Jan
; APPLICANT: Pastorekova, Silvia
; APPLICANT: Pastorek, Jaromir
; TITLE OF INVENTION: MN Gene and Protein
; NUMBER OF SEQUENCES: 96
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leona L. Lauder
; STREET: 369 Pine Street, Suite 610
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/787,739
; FILING DATE: 24-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,049
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/486,756
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/477,504
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/481,658
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,862
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,863
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/487,077
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauder, Leona L.
; REGISTRATION NUMBER: 30,863
; REFERENCE/DOCKET NUMBER: D-0021.4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-981-2034
; TELEFAX: 415-981-0332
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
;
; US-08-787-739-65

```

```

Query Match      0.3%; Score 76.2; DB 5; Length 105;
Best Local Similarity 82.9%; Pred. No. 5.6e-07;
Matches 87; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```

```

QY 10422 TTTTGTATTTTATAGAGAGGGGTTTCACCATGTTAGCCAGAGTGTCTGCATCTC 10481
      ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1 TTTTTCATCTTTAGTAGAGAGGGTTTCACCATATTGGCCAGGCTGCTCCAACATC 60
QY 10482 CTGACCTCGTATCGGCCACCTGAGCCTCCCAAAGTCTGGGAT 10576
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 61 CTGACCTGTGTATCCACACAGCTCGGCTCCCAAAGTCTGGGAT 105

```

```

RESULT 6
US-08-454-557C-91
; Sequence 91, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentlin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/454,557C
 FILING DATE: 30-MAY-1995
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Ludwig, Steven R.
 REGISTRATION NUMBER: 36, 203
 REFERENCE/DOCKET NUMBER: 0609, 3840003
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 91:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 84 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: both
 TOPOLOGY: both
 US-08-454-557C-91

Query Match	0.3%	Score 67;	DB 3;	Length 84;
Best Local Similarity	88.0%	Pred. No. 3.6e-05;		
Matches 73;	Conservative	0;	Mismatches 10;	Indels 0;
Gaps				
OY 10454	CCATTTCAGCCGAGNAGGCTGCACTCTCTGACACGCGATGACCGGCACCTAGACTCC	10513		
Db 1	CCATTTCATTCAGCGCGGTGTGAATCTCTTGACCTGTGATCCGCCCGCTCAGCTCC	60		
OY 10514	AAAGTCTGGGATTCACAGGTGTG	10536		
Db 61	AAAGTCTGGGATTCACACCGTGTG	83		

RESULT 7
 US-08-340-426D-91
 Sequence 91 Application US/08340426D
 Patent No. 5948634
 GENERAL INFORMATION:
 APPLICANT: de la Monte, Suzanne
 APPLICANT: Wands, Jack R.
 TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
 TITLE OF INVENTION: of Alzheimer's Disease
 NUMBER OF SEQUENCES: 121
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
 STREET: 1100 New York Avenue, Suite 600
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005-3934
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/340,426D
 FILING DATE: 14-NOV-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Ludwig, Steven R.
 REGISTRATION NUMBER: 36,203
 REFERENCE/DOCKET NUMBER: 0609,3840002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2600
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 91:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 84 base pairs

```

; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-340-426D-91

```

Query Match	0.3%;	Score 67;	DB 4;	Length 84;
Best Local Similarity	88.0%;	Pred. No.	3.6e-05;	
Matches 73;	Conservative	0;	Mismatches 10;	Indels 0;
				Gaps 0;

```

QY 10454 CCATCTTAGCAGGATGGTCTCATCTCTGTGACCTCGTATCCGCCACCTGAGCCCTCC 10513
      ||||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1 CCATGTTCAACAGGCTGGGTGTGCACTCTGACCTCGTATCCGCCGCCCTCAGCCCTCC 60

```

Qy 10514 AAAGTGTGGGATTACAGGTGTG 10536
|||||
Db 61 AAAGTGTGGGATTACAGCGTGT 83

RESULT 8
US-08-450-

; Sequence 91, Application US/08450673Q
; Patent No. 5948888

GENERAL INFORMATION:

APPLICANT: de la Monte, Suzanne

APPLICANT: Wallis, Jack R.
TITLE OF INVENTION: Neural Three

TITLE OF INVENTION: of Alzheimer's Disease

NUMBER OF SEQUENCES: 121

; CORRESPONDENCE ADDRESS:

ADDRESS: Sterne, Kessler, Goldstein &
CMBP: 1100 New York Avenue Suite 600

STREET: 1100 NEW
CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

; ZIP: 20005-3934

COMPUTER READABLE FORM:

```

; MEDIUM TYPE: floppy disk
;
; COMPUTER: IBM PC compat

```

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

```

; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/0

FILED DATE: 30-MAY-1993
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
CLASSIFICATION: 330

NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203

REFERENCE/DOCKET NUMBER: 06

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 371-2600
TELETYPE: (202) 371-3540

TELEFAX: (202) 374-2340
 INFORMATION FOR SEN ID NO: 91:

```

: INFORMATION FOR DEB ID NO: 21:
: SEQUENCE CHARACTERISTICS:

```

LENGTH: 84 base pairs

TYPE: nucleic acid

STRANDEDNESS: both

10P0LOGY: both
; 75-08-450-6735-01

05-08-430-0136-91

Query Match	0.38; Score 67; DB 4; Length 84;
-------------	----------------------------------

Best local similarity	88.0%;	Pred. No. 3.6e-05;			
Matches 73;	Conservative	0;	Mismatches 10;	Indels 0;	Gaps 0;

QY 10454 CCATGTTAGCCAGATGGTCTCGATCTCTGACCTCGTATCCGCCACCTGAGCCTCC 10513

Db 1 CCATGTTCAATCAGGCTGGTGTGGAAGCTCTGACCTCGTGATCCGCCCGCCTCAGGCTCCC 60

QY 10514 AAGTCTGGGATTACAGGTGTG 10536

Db 61 AAAGTGGTGGATTACACAGCGTG 83

```
RESULT 9
PCT-US95-17111A-91
; Sequence 91, Application PC/TUS9517111A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/17111A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/340,426
; FILING DATE: 14-NOV-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
PCT-US95-17111A-91

Query Match 0.3%; Score 67; DB 6; Length 84;
Best Local Similarity 88.0%; Pred. No. 3.6e-05;
Matches 73; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 10454 CCAATGTAGCAGGATGCTCGATCCGACCTCGATGCCGACCGACCTGACCTCC 10513
DB 1 CCAATGTATCAGGCTGCTGTCGAACTCTGATCGCTGATCCGCCCTCAGCTCC 60
QY 10514 AAAGTCTGGGATTATACAGTGTG 10536
DB 61 AAAGTCTGGGATTATCAAGCGTG 83

RESULT 10
US-08-454-557C-91/C
; Sequence 91, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.

COUNTRY: U.S.A.
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,557C
FILING DATE: 30-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.
REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 84 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: both
US-08-454-557C-91

Query Match 0.3%; Score 60.6; DB 3; Length 84;
Best Local Similarity 83.1%; Pred. No. 0.0007;
Matches 69; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 4867 CAGCCGTGATATTCAGACACTTGGAGAGCCGAGCAGGACGATCAGAGTCAATG 4926
DB 83 CAGCCTGTAATCCACACACTTGGAGGCTGAGGCGGCGATCAGAGTCAAGATT 24
QY 4927 CAAAGCCAGCCTGACCAATATGG 4949
DB 23 CGACACCAAGCCTGATGAACATGG 1

RESULT 11
US-08-340-426D-91/C
; Sequence 91, Application US/08340426D
; Patent No. 5948634
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; APPLICANT: Wands, Jack R.
; TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,426D
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ludwig, Steven R.
; REGISTRATION NUMBER: 36,203
; REFERENCE/DOCKET NUMBER: 0609.3840002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
```

; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-340-426D-91

Query Match 0.3%; Score 60.6; DB 4; Length 84;
Best Local Similarity 83.1%; Pred. No. 0.0007;
Matches 69; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 4867 CATGCTGTAAATTCAGACCTTTGGAGGCGGAGCGAGCATCATGAGTCAGGAATG 4926
DB 83 CACGCTTGTAAATCCAGACCTTTGGAGGCGGAGCGGAGCATCATGAGTCAGGAATG 24

QY 4927 CAAGACACGCTGACCAATATGG 4949
DB 23 CGACACACGCTGATGACATGG 1

RESULT 12
US-08-450-673C-91/C
; Sequence 91, Application US/08450673C
; Patent No. 5948888
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
US-08-450-673C-91

Query Match 0.3%; Score 60.6; DB 4; Length 84;
Best Local Similarity 83.1%; Pred. No. 0.0007;
Matches 69; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 4867 CATGCTGTAAATTCAGACCTTTGGAGGCGGAGCGAGCATCATGAGTCAGGAATG 4926
DB 83 CACGCTTGTAAATCCAGACCTTTGGAGGCGGAGCGGAGCATCATGAGTCAGGAATG 24
QY 4927 CAAGACACGCTGACCAATATGG 4949

DB 23 CGACACACGCTGATGACATGG 1

RESULT 13
PCT-US95-17111A-91/C
; Sequence 91, Application PC/TUS9517111A
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and
; TITLE OF INVENTION: Detection of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-3934
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 84 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
PCT-US95-17111A-91

Query Match 0.3%; Score 60.6; DB 6; Length 84;
Best Local Similarity 83.1%; Pred. No. 0.0007;
Matches 69; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY 4867 CATGCTGTAAATTCAGACCTTTGGAGGCGGAGCGAGCAATCATGAGTCAGGAATG 4926
DB 83 CACGCTTGTAAATCCAGACCTTTGGAGGCGGAGCGGAGCATCATGAGTCAGGAATG 24
QY 4927 CAAGACACGCTGACCAATATGG 4949
DB 23 CGACACACGCTGATGACATGG 1

RESULT 14
US-08-454-557C-70
; Sequence 70, Application US/08454557C
; Patent No. 5830670
; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne
; TITLE OF INVENTION: Neutral Thread Protein Gene Expression and Detection
; TITLE OF INVENTION: of Alzheimer's Disease
; NUMBER OF SEQUENCES: 121
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.

STREET 1100 New York Avenue, Suite 600
CITY Washington
STATE D.C.
COUNTRY U.S.A.
ZIP 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE Floppy disk
COMPUTER IBM PC compatible
OPERATING SYSTEM PC-DOS/MS-DOS
SOFTWARE PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER US/08/454,557C
FILING DATE 30-MAY-1995
CLASSIFICATION 514
ATTORNEY/AGENT INFORMATION:
NAME Ludwig, Steven R.
REGISTRATION NUMBER 36,203
REFERENCE/DOCKET NUMBER 0609,3840003
TELECOMMUNICATION INFORMATION:
TELEPHONE (202) 371-2600
TELEFAX (202) 371-2540
INFORMATION FOR SEQ ID NO: 70:
SEQUENCE CHARACTERISTICS:
LENGTH 78 base pairs
TYPE nucleic acid
STRANDEDNESS both
TOPOLOGY both
US-08-454-557C-70

Query Match	0.3%	Score 58.8	DB 3	Length 78
Best Local Similarity	84.6%	Pred. No.	0.0016	
Matches	66	Conservative	0	Mismatches 12
				Indels 0
				Gaps 0

Oy 14672 ACCACGCCCTAATTTTGTATTCTCGGTAGACGCGGTTTCACCAT3TTGGCCAG 14731
 ||||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 1 ACAACGCCCAGCTAATTTGATTCTTTAGTAGAGATGGCGTTCTCCCAT3TTCATCAG 60

```

Oy  14732 CTGGTCTGGAAGCTCTGA 14749
      ||||| ||||| |||||
Db   61 CTGGTCTGGAAGCTCTGA 78

```

RESULT 15
US-08-340-426D-70

Sequence 70, Application US/08340426D
; Patent No. 5948634

```

; GENERAL INFORMATION:
; APPLICANT: de la Monte, Suzanne

```

APPLICANT: Wands, Jack R.
TITLE OF INVENTION: Neural Thread Protein Gene Expression and Detection

```

; TITLE OF INVENTION:  Of Alzheimer's Disease
; NUMBER OF SEQUENCES:  121
;

```

```

;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
;

```

STREET: 1100 New
CITY: Washington

STATE: D.C.
COUNTRY: U.S.A.

ZIP: 20005-3934
COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: DOS 3.31

```

```

;
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
;

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/340,426D
;

```

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Ludwig, Steven R.

REGISTRATION NUMBER: 36,203
REFERENCE/DOCKET NUMBER: 0609.3840002

```

? TELECOMMUNICATION INFORMATION
? TELEPHONE: (202) 371-2600
?
? TELEFAX: (202) 371-2640
?
? INFORMATION FOR SEQ ID NO: 70:
?
? SEQUENCE CHARACTERISTICS:
?
? LENGTH: 78 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: both
?
? TOPOLOGY: both
US-08-340-426D-70

```

Query Match	0.3%	Score 58	DB 4	Length 78
Best Local Similarity	84.6%	Pred. No.	0.0016	
Matches	66	Conservative	0	Mismatches 12
				Indels 0
				Gaps 0

```

QY 14672 ACCACGCGCTAAATTTTGTATTTCTGCTAGACACGGGTTTCACCATGTGGCCAGG 14731
      || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 ACAACGCCACAGCTAAATTTGTATTTTGTAGAGATGGGGTTTCCCATGTCATCAGG 60

```

Qy	14732	CTGGTCTCGAACTCCTGA	14749
Db	61	CTGGTCTCGAACTCCTGA	78

```
Search completed: June 17, 2000, 17:39:10
Job time: 276280 sec
```

